2017

Distributed Kalman Filters over Wireless Sensor Networks: Data Fusion, Consensus, and Time-Varying Topologies

Jianming Zhou
Louisiana State University and Agricultural and Mechanical College

DISTRIBUTED KALMAN FILTERS OVER WIRELESS SENSOR NETWORKS: DATA FUSION, CONSENSUS, AND TIME-VARYING TOPOLOGIES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Division of Electrical and Computer Engineering

by
Jianming Zhou
B.S., Shanghai Jiao Tong University, China, 2013
M.S., Louisiana State University, USA, 2015
August 2017

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor and committee chair, Dr. Guoxiang Gu, for his helpful academic suggestions and patient guidance throughout my research and the preparation of this dissertation. His technical advice and support helped me overcome obstacles, kept me motivated, and made this research a meaningful learning experience.

Meanwhile, I would like to express my gratitude to Dr. Kemin Zhou, Dr. Morteza Naraghi-Pour, Dr. Mehdi Farasat, and Dr. Marcio de Queiroz for taking time out of their busy schedule and consenting to be a part of my committee, and for their valuable feedback.

In addition, I would like to thank the faculty, staff, and students in the Division of Electrical and Computer Engineering for all the help and support I have received.

I would like to deeply thank my dear parents, who sacrifice so much to make my dream come true. Thanks also go to my family, who have always been by my side to help me go through all the difficulty during the past four years.

Finally, I would like to thank Dr. Bixiang Tang, Dr. Abhishek Pandey, and Hyundeok Kang for their generous academic and moral support throughout my study at LSU and my life in the United States.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Kalman filtering is a widely used recursive algorithm for optimal state estimation of linear stochastic dynamic systems. The recent advances of wireless sensor networks (WSNs) provide the technology to monitor and control physical processes with a high degree of temporal and spatial granularity. Several important problems concerning Kalman filtering over WSNs are addressed in this dissertation. First we study data fusion Kalman filtering for discrete-time linear time-invariant (LTI) systems over WSNs, assuming the existence of a data fusion center that receives observations from distributed sensor nodes and estimates the state of the target system in the presence of data packet drops. Following [77], we focus on the single sensor node case and show that the critical data arrival rate of the Bernoulli channel can be computed by solving a simple linear matrix inequality problem. Then a more general scenario is considered where multiple sensor nodes are employed. We derive the stationary Kalman filter that minimizes the average error variance under a TCP-like protocol. The stability margin is adopted to tackle the stability issue. Second we study distributed Kalman filtering for LTI systems over WSNs, where each sensor node is required to locally estimate the state in a collaborative manner with its neighbors in the presence of data packet drops. The stationary distributed Kalman filter (DKF) that minimizes the local average error variance is derived. Building on the stationary DKF, we propose Kalman consensus filter for the consensus of different local estimates. The upper bound for the consensus coefficient is computed to ensure the mean square stability of the error dynamics. Finally we focus on time-varying topology. The solution to state consensus control for discrete-time homogeneous multi-agent systems over deterministic time-varying feedback topology is provided, generalizing the existing results. Then we study distributed state estimation over WSNs with time-varying communication topology. Under the uniform observability, each sensor node can closely track the dynamic state by using only its own observation, plus information exchanged with its neighbors, and carrying out local computation.

CHAPTER 1
INTRODUCTION

This chapter introduces the significance, applications and existing research works of Kalman filtering over wireless sensor networks (WSNs). The contribution and organization of this dissertation are discussed, followed by the introduction of notations.

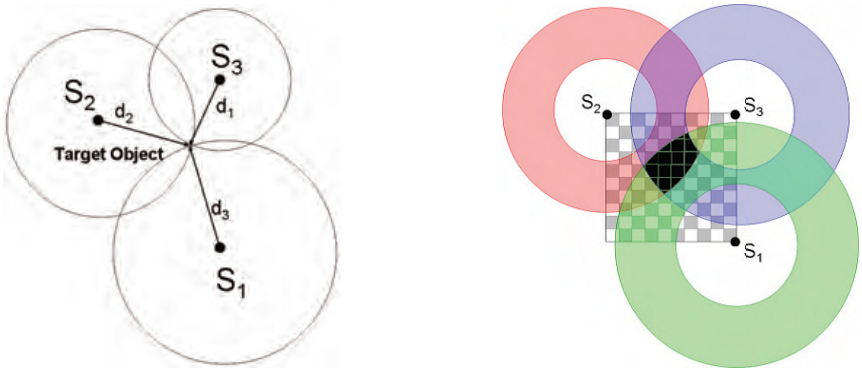
1.1 Motivation

Recent advances in micro-electro-mechanical systems, embedded microprocessor technology, and wireless communications facilitate the massive production of cheap, low-power, and long-lasting sensors of miniature sizes, integrated with the function of sensing, data computation and processing, and wireless communications. A wireless sensor network usually consists of a mesh of such sensor devices that are spatially distributed over an area of interest for specific monitoring tasks. The fast and remarkable development of WSNs provides the technology to monitor and control physical processes with a high degree of temporal and spatial granularity. Indeed, individual sensors in a WSN can sense and process data locally in real-time, provide resultant information concerning the observed events or processes, and communicate with data fusion centers or other sensor nodes in the network for collaboration tasks. In addition, the distributed nature of WSNs implies that the chance of a large number of these hardware units failing simultaneously is quite low, leading to improved robustness compared to traditional monitoring systems. Applications of WSNs are ubiquitous, ranging from environmental surveillance [15] and biomedical health monitoring [24, 58], to mobile sensing [52] and vehicle navigation and control [48, 75], along with many others [44].

Dynamic state estimation has its great importance in various applications such as detection, tracking, and control. For discrete-time linear stochastic dynamic systems, Kalman filtering [39, 4] is one of the most widely used recursive algorithms, which computes the state estimate in the sense of minimum mean square error (MMSE) when the process and measurement noises are Gaussian distributed.

Implementing Kalman filtering with WSN technology has received increased and great attention in recent research and literatures. One of the most influential and important applications of Kalman filtering over WSNs is the localization and tracking of static or mobile targets [76, 40, 97, 25, 65, 89], which is usually considered as a fundamental requirement in other more sophisticated location-aware applications of WSNs, such as indoor positioning, vehicle navigation, military security, etc.

Triangulation and trilateration [76] are two widely used techniques for localization and tracking. The former uses distances, angles and trigonometric relationships to identify the position of the object of interest, while the latter only relies on the distance measurements to locate the target. The scenario for trilateration is illustrated in Figure 1.1a, where d_1, d_2, and d_3 are the distances from the fixed sensor nodes S_1, S_2, and S_3 with known locations to the mobile target node, respectively. Ideally the mobile node can be located at the intersection of the three circles. However, in the presence of measurement noise, the distance measurement can fluctuate within the margin of the ring. The task of tracking thus becomes more difficult since the mobile node can be located anywhere in the dark overlapped region in Figure 1.1b.



(a) Trilateration (b) Trilateration with noise

Figure 1.1: [76] Trilateration Localization

Commonly used types of range (distance) measurements in WSNs include angle of arrival (AOA) [19, 73], time of arrival (TOA) [51, 61], time difference of arrival (TDOA) [35, 40, 59], and received signal strength (RSS) [60, 61, 97]. Typically for TDOA, the transmitter known as beacon disseminates information in a radio frequency (RF) signal, together with

a concurrent ultrasound pulse. The receiver known as listener uses the time difference of arrival of the two signals to compute its distance from the beacon by taking advantage of the speed difference between ultrasound pulse (speed of sound) and RF signal (speed of light). The RSS, measured in dB, typically depends on three key factors: path loss, shadow fading, and fast fading, i.e., [82, 84, 97]

$$p = \kappa - 10\gamma \log(d) + \psi,$$

where p is the RSS, κ is a constant determined by the transmitted power, wavelength, antenna height, etc., γ denotes the slope index, d represents the distance between the signal transmitter and the signal receiver, ψ is a Gaussian variable with mean zero.

Existing indoor location systems integrated with WSNs can be grouped into two classes based on their architectures. For active mobile architecture, the mobile target is equipped with an active transmitter that broadcasts a message, such as an RF signal coupled with an ultrasound pulse, to all the sensor nodes in the WSN. The receiver deployed at each sensor node extracts range measurement from the broadcast and sends this distance information to a data fusion center for real-time tracking of the mobile target. Examples of this architecture include the Active Badge system [90], the Bat system [34], along with several others. In contrast, for passive mobile architecture, beacons deployed at fixed sensor nodes send their location information to the moving node. The moving node then uses the set of range measurements to estimate its own position in real-time. The Cricket system [63, 78] is a widely used example that follows the passive mobile architecture.

Different dynamic models can be adopted for the mobile target based on its motion pattern to achieve optimal tracking performance. The Position (P) model works well when the position is mostly constant, and the velocity can be treated as noise. The Position-Velocity (PV) model can be chosen when the velocity remains unchanged most of the time,

and the acceleration can be regarded as noise. When the acceleration is mostly constant, the Position-Velocity-Acceleration (PVA) model tends to provide the best tracking results.

Now we focus on the Cricket system to show how Kalman filtering over WSN can be applied for indoor localization and tracking. Consider a Cricket system where each of the N sensor nodes transmits its location in a RF signal with a concurrent ultrasound pulse to the mobile node that then gathers a set of range measurements based on TDOA to estimate its own location. We assume that the mobile node moves in a two dimensional space following the PVA model. Let the state vector be specified by

$$\mathbf{x}(k) = [x(k)\ y(k)\ \dot{x}(k)\ \dot{y}(k)\ \ddot{x}(k)\ \ddot{y}(k)]',$$

where $x(k)$ and $y(k)$ respectively denotes the position of the mobile node in the x and y directions at time stamp k. It follows that

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + \mathbf{w}(k),$$

where process noise $\mathbf{w}(k)$ follows Gaussian distribution with mean zero and covariance $Q(k) \geq 0$, T_s is the time-step for discretization (the time interval between previous time instant and current time instant), and

$$A = \begin{bmatrix} I_2 & T_s I_2 & \frac{1}{2}T_s^2 I_2 \\ 0 & I_2 & T_s I_2 \\ 0 & 0 & I_2 \end{bmatrix}.$$

It is noted that the real distance from the ith sensor node located at position (coordinates) (x_i, y_i) to the mobile node at time k is given by

$$d_i[\mathbf{x}(k)] = \sqrt{[x(k) - x_i]^2 + [y(k) - y_i]^2}.$$

Let observation vector $\mathbf{z}(k)$ be the collection of all the N distance measurements. It follows that

$$\mathbf{z}(k) = h[\mathbf{x}(k)] + \mathbf{v}(k), \tag{1.1}$$

where measurement noise $\mathbf{v}(k)$ follows Gaussian distribution with mean zero and covariance $R(k) > 0$, and

$$h[\mathbf{x}(k)] = \begin{bmatrix} d_1[\mathbf{x}(k)] & \dots & d_N[\mathbf{x}(k)] \end{bmatrix}'.$$

Kalman filtering generally provides a decent solution for tracking over WSNs since it is capable of filtering out the process and measurement noises with low computation complexity and memory requirement. However the measurement model in (1.1) is a nonlinear function of the state vector, which implies that standard Kalman filtering cannot be applied in this case since it is specifically proposed for linear systems. On the other hand, the extended Kalman filtering (EKF) [76, 40] has been proposed and studied for state estimation of nonlinear systems in recent years. EKF is derived from the standard Kalman filtering by linearizing the nonlinear state and measurement model around the latest state estimate and the predicted state, respectively. Let $\hat{\mathbf{x}}(k)$ be the prediction of $\mathbf{x}(k)$ at time $k-1$, and $\Sigma(k)$ be the corresponding error covariance. Similarly let $\hat{\mathbf{x}}(k|k)$ be the estimation of $\mathbf{x}(k)$ at time k, and $\Sigma(k|k)$ be the corresponding error covariance. The measurement model in (1.1) can be linearized as

$$\mathbf{z}(k) = H(k)\mathbf{x}(k) + \mathbf{v}(k)$$

by taking the Jacobian matrix $H(k)$ as

$$H(k) = \frac{\partial h[\mathbf{x}(k)]}{\partial \mathbf{x}(k)}\Big|_{\mathbf{x}(k)=\hat{\mathbf{x}}(k)} = \begin{bmatrix} \frac{\hat{x}(k)-x_1}{d_1[\mathbf{x}(k)]} & \frac{\hat{y}(k)-y_1}{d_1[\mathbf{x}(k)]} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\hat{x}(k)-x_N}{d_N[\mathbf{x}(k)]} & \frac{\hat{y}(k)-y_N}{d_N[\mathbf{x}(k)]} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then the time update phase of the EKF algorithm, which provides the estimated state for

the next time instant, is specified by

$$\hat{\mathbf{x}}(k) = A\hat{\mathbf{x}}(k-1|k-1),$$
$$\Sigma(k) = A\Sigma(k-1|k-1)A' + Q(k-1);$$

the measurement update phase of EKF, which corrects the estimated state by utilizing the observation vector, is specified by

$$K(k) = \Sigma(k)H(k)'\left[R(k) + H(k)\Sigma(k)H(k)'\right]^{-1},$$
$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k) + K(k)\left\{\mathbf{z}(k) - h[\hat{\mathbf{x}}(k)]\right\},$$
$$\Sigma(k|k) = \Sigma(k) - K(k)H(k)\Sigma(k).$$

The above equations can be evaluated iteratively to track the moving node.

Next we consider a different example in [16], which deals with the tracking of a projectile with linear state and measurement models. Let the state vector be specified by

$$\mathbf{x}(k) = \left[\dot{x}(k)\ \dot{y}(k)\ \dot{z}(k)\ x(k)\ y(k)\ z(k)\right]',$$

where $x(k)$, $y(k)$, and $z(k)$ denote the position of the projectile in the three spatial dimensions respectively, with $z(k)$ being the vertical one. The motion of the projectile is described by state dynamics

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + b + \mathbf{w}(k)$$

with $\mathbf{w}(k)$ being the process noise, g being the gravity constant, T_s being the discretization time-step,

$$A = \begin{bmatrix} I_3 & 0 \\ T_s I_3 & I_3 \end{bmatrix}, \quad b = \begin{bmatrix} 0 & 0 & -gT_s & 0 & 0 & -\frac{1}{2}gT_s^2 \end{bmatrix}'.$$

A WSN of N sensor nodes is employed to measure and estimate the position of the projectile.

It is assumed that each node measures the position in either (i) the two horizontal dimensions, or (ii) the combination of one horizontal dimension and the vertical dimension. Hence the measurement $\mathbf{z}_i(k)$ at the ith sensor node follows

$$\mathbf{z}_i(k) = C_i\mathbf{x}(k) + \mathbf{v}_i(k),$$

where $C = \begin{bmatrix} 0 & \text{diag}(1,1,0) \end{bmatrix}$, which corresponds to case (i), or $C = \begin{bmatrix} 0 & \text{diag}(1,0,1) \end{bmatrix}$, which corresponds to case (ii), and $\mathbf{v}_i(k)$ is the measurement noise at node i. In [16], distributed Kalman filtering algorithms based on diffusion strategies are proposed for the WSN, so that each sensor node is able to locally estimate the state of the projectile in a collaborative manner, in spite of the restriction that individual nodes do not have direct position measurements in all the three dimensions. This numerical example will be modified and used for simulation in Chapter 3.

Although WSN presents attractive and compelling features, challenges associated with its implementation have to be addressed. Implementing Kalman filtering with the WSN technology normally requires the usage of wireless communication channels for data transmission. However, due to the low power nature and requirement of long-lasting deployment, communications between adjacent embedded sensors or between embedded sensors and data fusion center are often affected by range and other environmental elements, inducing the frequent presence of data drops and communication delays. This contrasts to the traditional ways to implement Kalman filtering, and inevitably results in the deterioration of estimation performance. Consequently, there has been a major development in the study of Kalman filtering in the presence of data packet drops [77, 62, 46, 74, 95, 37, 94, 96, 47], distributed Kalman filtering [53, 54, 55, 1, 16, 79, 41, 70, 88], distributed Kalman filtering with intermittent observations [43], etc. Inspired by the aforementioned applications and challenges, we are set to study several fundamental issues involved in Kalman filtering over WSN.

1.2 Literature Survey

One important feature of Kalman filtering over WSNs is that measurement acquisition and data processing may take place in different geographical locations, which motivates the study of Kalman filtering schemes where intermittent observations have to be dealt with. The basic problem of Kalman filtering in the presence of data packet drops was first formulated and studied in [77], where the measurements (observations) obtained by a single sensor node are transmitted over a data packet drop channel to the fusion center that then computes the state estimate. Since then, many researchers have studied various aspects of the problem by imposing different assumptions on communication channel model and system structure. Channel model describes the stochastic properties of the observation dropouts.

- Many papers consider the channel distortion induced by data packet drops as an independent and identically distributed (i.i.d.) Bernoulli random process. It is shown in [77] that there exists an infimum for the data arrival probability of the communication channel, referred to as critical arrival rate, below which the estimation error covariance becomes unbounded. The lower and upper bounds of the critical arrival rate are derived, and the two bounds coincide with each other when observation matrix C in the measurement model is invertible. This condition was relaxed in [62] to only requiring C being invertible on the observable subspace. Nevertheless these two invertibility conditions are hard to satisfy in practice. In [46], the authors show that the critical arrival rate is a function of system matrix A and observation matrix C, independent of the noise statistics and initial conditions. It is further claimed that the lower bound in [77] is indeed the critical value if (C, A) is detectable and the unstable eigenvalues of matrix A have distinct absolute values. This conclusion turns out to be incorrect based on the results in [14] and [7], which imply that for systems whose measurement has dimension one (scalar), the critical rate can be expressed in a closed form as a function of the Mahler measure of matrix A. More details concerning this result will be presented in Subsection 2.1.3.

- Another channel model extensively used in existing literatures is the Gilbert-Elliott model [26, 23], which was first introduced into the context of Kalman filtering with intermittent observations in [37]. Sufficient condition for the peak covariance stability is derived in [37], and turns out to be necessary for scalar systems. A simpler and less conservative sufficient condition for the stability of the peak covariance process is provided in [94]. In [96], necessary and sufficient conditions for the peak covariance stability are derived for second-order systems as well as higher-order systems of certain classes. In [47], a class of non-degenerate systems is proposed, and the corresponding stability results are established. Other network models are also investigated [17, 72].

In spite of the considerable effort in finding the stability conditions for Kalman filtering in the presence of data packet drops, a complete answer is not available yet, even when the simple Bernoulli channel model is used. It remains an open problem to compute the exact value of the critical arrival rate for general dynamic systems where no restrictions are imposed on the structure of system matrix A and observation matrix C. This problem will be discussed and the corresponding solution will be provided in Chapter 2.

The recent advances in WSN technology also boost the use of multiple sensors for distributed Kalman filtering in various applications ranging from civil to military fields. Ideally, all the raw measurements from distributed wireless sensor nodes can be gathered at a data fusion center to compute the globally optimal state estimate. However this scheme may not be feasible in some practical cases due to the limited channel bandwidth, restricted power consumption, and significant communication delays. Therefore, many researchers start to investigate the distributed Kalman filter (DKF), where each sensor node in the WSN can compute local estimates via Kalman filtering based on its own observations and the information sent from its neighboring sensors located within a predefined transmission radius. Compared with Kalman filtering using fusion center, DKF improves the resilience of WSNs to isolated points of failure. Moreover, DKF can be quite useful if individual sensor nodes are required to execute multiple tasks and need real-time local estimates to make on-site deci-

sions. DKF algorithms need to be devised to ensure the stability of the associated estimation error dynamics.

- Several fundamental and enlightening approaches and strategies for distributed Kalman filtering are proposed by Olfati-Saber in [53, 54, 55]. In [53], the author addresses the DKF problem by reducing it to two separate consensus problems in terms of weighted measurements and inverse-covariance matrices, which are then solved in a distributed way using low-pass and band-pass consensus filters respectively. The resulting DKF algorithm is only applicable to sensor nodes with identical observation matrix, which requires the dynamic system to be observable by all the nodes in WSN. This DKF algorithm is modified in [54], where two identical high-pass consensus filters are used for the fusion of sensor measurements and covariance information so that the DKF allows sensor nodes to have different observation matrices. A continuous-time distributed Kalman-Bucy filter is introduced in [54] as well, giving rise to a new DKF algorithm based on consensus of estimates. A common defect of these two algorithms is that both are derived from the discretization of continuous-time consensus filters. In fact consensus achieved by continuous-time filters may not hold after discretization. Yet the author fails to validate the stability in a rigorous way. In [55], Kalman consensus filter (KCF) is designed with the objective that each sensor node can locally estimate the state of the target, and reach consensus on state estimate. The author points out the computational unscalability of the optimal KCF, and proposes a scalable suboptimal KCF with formal stability and performance analysis.

- Various results concerning DKF are presented in other works. The DKF algorithm proposed in [1] is based on the standard Kalman filtering with one extended step where sensor nodes merge their estimates by a weighted average approach. The weights are optimized to yield a small estimation error covariance in the steady-state. Only state estimates are exchanged among neighboring nodes so that the bandwidth requirements are reduced. In [16], distributed Kalman filtering, fixed-lag smoothing and fixed-point smoothing are studied, and diffusion strategies are adopted to address these problems.

Well-defined expressions are provided for the steady-state mean square performance of the proposed algorithms. In [79], the authors investigate the globally optimal distributed Kalman filtering fusion where the covariance matrices of measurement noise and estimation error are singular. The problem of distributed state estimation for linear time-varying systems with intermittent observations is considered in [43]. An optimal KCF is derived by minimizing the mean square estimation error at each sensor node, and a suboptimal filter is proposed for scalability considerations.

1.3 Dissertation Contribution and Organization

Several fundamental problems concerning Kalman filtering over WSNs are addressed in this dissertation. Chapter 1 briefly discusses the background knowledge and applications of Kalman filtering over WSN, followed by the literature survey that motivates our research.

The main results and contributions are presented in Chapters $2 \sim 4$. In Chapter 2, we study data fusion Kalman filtering for discrete-time linear time-invariant (LTI) systems over WSN, assuming the existence of a data fusion center that receives measurements from embedded sensor node(s) and estimates the state of the target system in the presence of data packet drops. First we consider the case discussed in [77], where only one sensor node is employed to obtain and transmit observations in a single packet to the data fusion center. It is shown that the widely studied critical arrival rate of the Bernoulli channel can be computed by solving a set of linear matrix inequalities (LMIs). Then a more generalized setting is considered, where multiple sensor nodes transmit their measurements to the fusion center through different communication channels. This scenario makes the problem harder to analyze since the aforementioned arguments on critical arrival rates of individual channels are no longer applicable. We derive the stationary Kalman filter that minimizes the average error variance in the steady-state at the data fusion center, based on the stabilizing solution to a modified algebraic Riccati equation (MARE). The stability margin, which can be computed

by solving an LMI problem, is adopted to address the stability issue. It can serve as an important reference to decide the number and disposition of the sensor nodes.

Although both distributed Kalman filtering and the effect of packet drops have been widely studied in the past decades, only a few literatures consider the occurrence of data packet drops in the design of DKF. Therefore, in Chapter 3, we investigate distributed Kalman filtering over WSN, where each sensor node is required to locally estimate the state of a discrete-time LTI system using its own observations and those transmitted from its neighbors in the presence of data packet drops. This is an optimal one-step prediction problem under the framework of distributed estimation, assuming the TCP-like protocol [74, 77]. We first examine a general MMSE estimation problem, and apply the results to derive the stationary DKF that minimizes the local average error variance in the steady-state at each sensor node. The optimal estimation gain is presented in terms of the stabilizing solution to the corresponding MARE. The stability issue is addressed by adopting the stability margin. Following [55], we also consider designing KCF in the presence of data packet drops for the consensus of different local estimates. The KCF, consisting of the stationary DKF and a consensus term of prior estimates, is proposed, followed by the stability analysis. It is shown that the proposed KCF outperforms the stationary DKF in general.

In many practical applications, there are situations that the set of sensor nodes employed for observation and state estimation changes as time proceeds. For example, some sensor nodes join the set when the moving target comes closer to their locations, while others drop out of the set as the target moves away. Sometimes the WSN follows a predefined pattern to execute the observation and tracking tasks from time to time for the purpose of energy saving and device maintenance. In these cases, the Bernoulli processes adopted in Chapters 2 and 3 are no longer suitable for the modeling of communication channels. Instead, the communication between sensor nodes can be encoded through a time-varying topology (graph). In Chapter 4, based on the review of several restrictive state consensus protocols in Xiao et al. (2005) [92] and Jadbabaie et al. (2003) [38], we provide the solution

to state consensus control for discrete-time homogeneous multi-agent systems (MASs) over deterministic time-varying feedback topology, which not only agrees with, but also generalizes the results in [92] and [38]. Then we focus on the distributed state estimation over the WSN with deterministic time-varying topology. In our proposed protocol, the observation about the target system is required to be available at only one or a few nodes at each time instant so that the communication overhead between the target system and individual sensor nodes can be lowered to a large extent. Under the uniform observability of the time-varying graph, each sensor can closely track the dynamic state only by using the observations if obtained, exchanging information with its neighbors, and carrying out local computation.

Chapter 5 concludes the whole dissertation and presents some ideas about possible research topics for future work.

1.4 Notations

The notations in this dissertation are fairly standard. The symbols I_n, and $\mathbf{1}_n$ stand for the identity matrix of dimension $n \times n$, and the column vector of dimension n with all components one, respectively. Transpose and conjugate transpose are denoted by "$'$" and the superscript of "$*$", respectively. The eigenvalue is denoted by $\lambda(\cdot)$. The spectral radius is denoted by $\rho(\cdot)$. The Kronecker product and Hadamard product are denoted by"$\otimes$" and "$\circ$" respectively. The expectation and covariance operations are denoted by $\mathrm{E}\{\cdot\}$ and $\mathrm{Cov}\{\cdot\}$ respectively. Finally we denote $\mathrm{diag}\{\cdot\}$ as the (block) diagonalization operation, $\mathrm{vec}\{\cdot\}$ as the vectorization operation, and $\mathrm{col}\{\cdot\}$ as the columnization operation. Other notations will be made clear as we proceed.

CHAPTER 2
DATA FUSION KALMAN FILTERING

In this chapter we study data fusion Kalman filtering over WSN, where the data fusion center is employed to estimate the state of a discrete-time dynamic system based on the observations transmitted from distributed sensor node(s) in the presence of data packet drops.

2.1 Single Data Packet Drop Channel

This section is focused on the case discussed in [77], where the measurement of the sensor node is transmitted to the fusion center through a single data packet drop channel.

2.1.1 Mean Square Stability

We start by introducing the mean square (MS) stability for the feedback system configured in Figure 2.1. The plant model is described by transfer matrix $G(z)$ with state space representation

$$G(z) = C_g(zI - A_g)^{-1}B_g,$$

where $A_g \in \mathbb{R}^{n\times n}$, $B_g \in \mathbb{R}^{n\times m}$, and $C_g \in \mathbb{R}^{m\times n}$. The following assumptions are made, which may be required in various situations.

Assumption 1. *A_g is a Schur stability matrix.*

Assumption 2. $\text{rank}\{B_g\} = \text{rank}\{C_g\} = m$.

Let the multiplicative noise (fading channel) in Figure 2.1 be specified by $\delta(k)I_m$ where $\{\delta(k)\}$ is a white random process with mean zero and variance σ^2, i.e.,

$$\text{E}\{\delta(k)\} = 0,\ \text{E}\{\delta(k)^2\} = \sigma^2 \ \ \forall\ k \geq 0.$$

The MS stability is defined next.

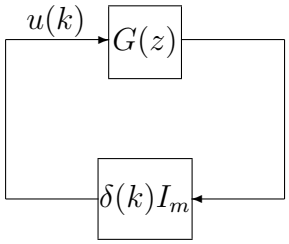


Figure 2.1: Feedback system with multiplicative noise over single feedback channel

Definition 1. [13] *Under Assumption* 1*, the closed-loop system in Figure* 2.1 *is said to be MS stable, if the variance of* $\{u(k)\}$ *is asymptotically bounded, i.e.,*

$$\lim_{k\to\infty} \mathrm{E}\{\|u(k)\|^2\} < \infty.$$

A well-known MS stability condition in [13] is presented in the next lemma.

Lemma 1. *Under Assumption* 1*, the following statements are equivalent:*

(i) *The feedback system in Figure* 2.1 *is MS stable;*

(ii) *There exists* $X_g > 0$ *satisfying the LMI:*

$$X_g > A_g X_g A_g' + \sigma^2 B_g C_g X_g C_g' B_g'; \tag{2.1}$$

(iii) *There exists* $Y_g > 0$ *satisfying the LMI:*

$$Y_g > A_g' Y_g A_g + \sigma^2 C_g' B_g' Y_g B_g C_g. \tag{2.2}$$

It is easy to see that (2.1) and (2.2) are dual to each other. The following result provides a closed-form test for the MS stability, which can be valuable in some applications.

Lemma 2. *Let* $H = (C_g \otimes C_g)(I - A_g \otimes A_g)^{-1}(B_g \otimes B_g)$*. Under Assumption* 1*, the feedback system in Figure* 2.1 *is MS stable, if and only if* $\rho(H) < \sigma^{-2}$*.*

Proof. Suppose that the feedback system in Figure 2.1 is MS stable. Then LMI (2.1) in Lemma 1 admits a solution $X_g > 0$ that has no loss of generality. Denote

$$Q_\varepsilon = X_g - A_g X_g A_g' - \sigma_\varepsilon^2 B_g C_g X_g C_g' B_g',$$

where $\sigma_\varepsilon^2 = \varepsilon\sigma^2$. Then $Q_\varepsilon > 0$ for $\forall\ \varepsilon \in (0,\ 1]$. Taking the vectorization operation on both sides of the above equation yields

$$\mathrm{vec}(Q_\varepsilon) = [I - A_g \otimes A_g - \sigma_\varepsilon^2 (B_g \otimes B_g)(C_g \otimes C_g)]\mathrm{vec}(X_g).$$

By Schur stability of A_g, there holds $\rho\,(A_g \otimes A_g) < 1$. Then

$$\rho\left\{(A_g \otimes A_g) + \sigma_\varepsilon^2 (B_g \otimes B_g)(C_g \otimes C_g)\right\} < 1\ \ \forall\ \varepsilon \in (0,\ 1]$$

in light of the continuity argument and the hypothesis on MS stability. The above inequality implies that

$$\det\left[I - A_g \otimes A_g - \sigma_\varepsilon^2 (B_g \otimes B_g)(C_g \otimes C_g)\right] \neq 0 \tag{2.3}$$

for $\forall\ \varepsilon \in (0,\ 1]$ with $\det(\cdot)$ denoting the determinant. Then we have

$$\det\left[I - \sigma_\varepsilon^2 (C_g \otimes C_g)(I - A_g \otimes A_g)^{-1}(B_g \otimes B_g)\right] \neq 0.$$

Therefore $\det(I - \varepsilon\sigma^2 H) \neq 0$ for $\forall\ \varepsilon \in (0,\ 1]$, concluding $\rho(H) < \sigma^{-2}$.

Conversely if $\rho(H) < \sigma^{-2}$, then $\rho(H) < (\varepsilon\sigma^2)^{-1} = \sigma_\varepsilon^{-2}$ for $\forall\ \varepsilon \in (0,\ 1]$. By the continuity argument, there exists a sufficiently small $\varepsilon_{\mathrm{m}} > 0$ such that the feedback system is MS stable for multiplicative white noise $\{\delta_\varepsilon(k)\}$ having variance $\sigma_{\varepsilon_{\mathrm{m}}}^2$. Hence there exists $X_{\varepsilon g} > 0$ satisfying LMI

$$X_{\varepsilon g} > A_g X_{\varepsilon g} A_g' + \sigma_\varepsilon^2 B_g C_g X_{\varepsilon g} C_g' B_g' \tag{2.4}$$

for $\forall\ \varepsilon \in (0,\ \varepsilon_{\mathrm{m}}]$. Recall that $\rho\,(A_g \otimes A_g) < 1$. Hence inequality (2.3) holds for $\forall\ \varepsilon \in (0,\ \varepsilon_{\mathrm{m}}]$

by making use of the previous arguments leading to (2.3). We claim that $\varepsilon_{\rm m} = 1$. If $\varepsilon_{\rm m} \neq 1$, then there exists some $\varepsilon_0 \in (0,\ 1]$ satisfying $\varepsilon_0 > \varepsilon_{\rm m}$ such that inequality (2.3) fails. That is,

$$\det\left[I - A_g \otimes A_g - \sigma_{\varepsilon_0}^2 (B_g \otimes B_g)(C_g \otimes C_g)\right] = 0.$$

It follows from previous derivations that $\det(I - \sigma_{\varepsilon_0}^2 H) = 0$, which contradicts to the hypothesis that $\varepsilon\rho(H) < \sigma^{-2}$ for $\forall\ \varepsilon \in (0,\ 1]$. Therefore $\varepsilon_0 \notin (0,\ 1]$ and LMI (2.4) admits a solution $X_{\varepsilon g} > 0$ for $\varepsilon = 1$, which reduces to (2.1),thereby concluding the MS stability. □

Lemma 2 indicates that $\sigma_{\rm sup} := 1/\sqrt{\rho(H)}$ is a critical value in the sense that MS stability holds for all $\sigma < \sigma_{\rm sup}$. While both Lemma 1 and Lemma 2 provide an elegant MS stability condition, neither one can be used in Kalman filtering synthesis directly, since the unknown estimation gain is involved in $G(z)$. One may follow the iterative algorithm in [22] to search for $\sigma_{\rm sup}$ and estimator parameters in alternation by fixing the other, but it becomes a nonlinear programming problem. Building on Lemma 1, we have the following MS stability result, which is fundamental and more useful to the proof of the main results of this section.

Lemma 3. *Under Assumptions* 1 *and* 2*, the following statements are equivalent:*

(i) *The feedback system in Figure* 2.1 *is MS stable;*

(ii) *There exists* $\Phi > 0$ *such that*

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_\Phi(e^{j\omega}) G_\Phi(e^{j\omega})^* d\omega\right) < \sigma^{-2}, \tag{2.5}$$

where $G_\Phi(z) = \Phi^{-1} G(z) \Phi$*;*

(iii) *There exists* $\Psi > 0$ *such that*

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_\Psi(e^{j\omega})^* G_\Psi(e^{j\omega}) d\omega\right) < \sigma^{-2}, \tag{2.6}$$

where $G_\Psi(z) = \Psi G(z) \Psi^{-1}$.

Proof. For (i) $\Rightarrow$ (iii): There exists $Y_g > 0$ satisfying LMI (2.2), if the feedback system in Figure 2.1 is MS stable, in light of Lemma 1. Substituting

$$Y_g = zA_g' Y_g + z^{-1} Y_g A_g - A_g' Y_g A_g + (z^{-1} I - A_g') Y_g \, (zI - A_g)$$

into LMI (2.2) for $|z| = 1$ with rearrangement yields

$$\begin{aligned} \sigma^2 C_g' B_g' Y_g B_g C_g &< zA_g' Y_g + z^{-1} Y_g A_g - 2A_g' Y_g A_g + (z^{-1} I - A_g') Y_g (zI - A_g) \\ &= (z^{-1} I - A_g') Y_g A_g + A_g' Y_g (zI - A_g) + (z^{-1} I - A_g') Y_g (zI - A_g). \end{aligned}$$

Multiplying $B_g'(z^{-1} I - A_g')^{-1}$ from left and $(zI - A_g)^{-1} B_g$ from right to the above inequality lead to

$$\sigma^2 G(z)^* B_g' Y_g B_g G(z) < B_g' Y_g A_g (zI - A_g)^{-1} B_g + B_g' (z^{-1} I - A_g')^{-1} A_g' Y_g B_g + B_g' Y_g B_g.$$

Computing average over the unit circle gives

$$B_g' Y_g B_g > \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} G(e^{j\omega})^* B_g' Y_g B_g G(e^{j\omega}) d\omega.$$

Denote $\Psi^2 = B_g' Y_g B_g > 0$. Then $\Psi > 0$ by $\text{rank}\{B_g\} = m$ and $Y_g > 0$. The above is equivalent to

$$I > \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} G_\Psi(e^{j\omega})^* G_\Psi(e^{j\omega}) d\omega,$$

that is in turn equivalent to (2.6).

For (iii) $\Rightarrow$ (i): By the Schur stability of A_g, there exists $Y_g \geq 0$ to the Lyapunov equation

$$Y_g = A_g' Y_g A_g + C_g' \Psi^2 C_g. \tag{2.7}$$

Note that the power spectral density $G_\Psi(e^{j\omega})^* G_\Psi(e^{j\omega})$ admits the decomposition

$$\begin{aligned} G_\Psi(e^{j\omega})^* G_\Psi(e^{j\omega}) &= \Psi^{-1} B_g' Y_g B_g \Psi^{-1} \\ &\quad + \Psi^{-1} B_g' Y_g A_g (e^{j\omega} I - A_g)^{-1} B_g \Psi^{-1} + \Psi^{-1} B_g' (e^{-j\omega} I - A_g')^{-1} A_g' Y_g B_g \Psi^{-1}. \end{aligned}$$

If (2.6) holds, then Y_g satisfies

$$\sigma^2 \Psi^{-1} B_g' Y_g B_g \Psi^{-1} < I_m. \tag{2.8}$$

Substituting (2.8) into (2.7) yields

$$Y_g > A_g' Y_g A_g + \sigma^2 C_g' \Psi (\Psi^{-1} B_g' Y_g B_g \Psi^{-1}) \Psi C_g = A_g' Y_g A_g + \sigma^2 C_g' B_g' Y_g B_g C_g,$$

which further implies $Y_g > 0$, thereby concluding the proof for (2.6). Since (2.5) is dual to (2.6), its proof is omitted. □

2.1.2 Stationary Data Fusion Kalman Filter

The problem of Kalman filtering with intermittent observations was first formulated and studied in Sinopoli et al (2004) [77]. Consider the networked system described by

$$x(k+1) = Ax(k) + w(k), \tag{2.9a}$$

$$y(k) = Cx(k) + v(k), \tag{2.9b}$$

where $x(k) \in \mathbb{R}^n$ denotes the state of the dynamic target system, and $y(k) \in \mathbb{R}^m$ denotes the measurement obtained by the sensor node at time index k. The process noise $w(k)$ and the measurement noise $v(k)$ are independent white processes with mean zero and covariance $Q \geq 0$ and $R > 0$, respectively, uncorrelated to the initial state $x(0) = x_0$ that has mean $\overline{x}_0$ and covariance Σ_0. By convention, it is assumed that x_0, $w(k)$, and $v(k)$ are jointly Gaussian

distributed. Since the measurement information is transmitted from the sensor node to the data fusion center over a packet drop channel, the signal at the receive end is given by

$$s(k) = \gamma(k)y(k) \tag{2.10}$$

instead of $y(k)$ itself, where $\gamma(k)$ is imposed to represent the network distortion induced by data packet drops. A simple Bernoulli model is adopted, assuming i.i.d. stationary process, for $\{\gamma(k)\}$ with probability

$$\mathrm{P}\{\gamma(k) = 1\} = p > 0 \ \ \forall\, k \geq 0.$$

Moreover, $\gamma(k)$ is independent of the noise and the initial state. Based on (2.10), observation $y(k)$ is received successfully by the fusion center at time k if $\gamma(k) = 1$; otherwise all the components of $y(k)$ are lost. The complete setup is depicted in Figure 2.2. The following assumption is made without loss of generality.

Assumption 3. $\mathrm{rank}\{C\} = m$.

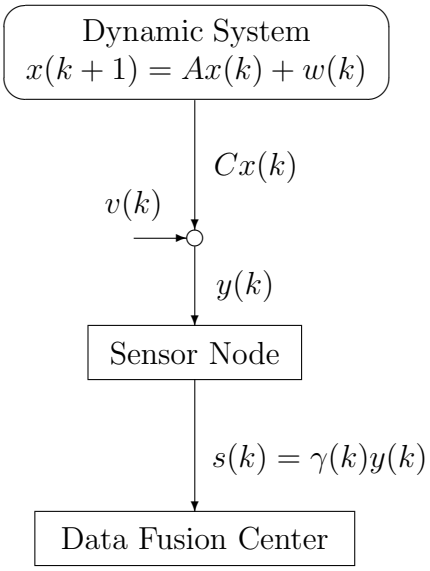


Figure 2.2: Setup for data fusion Kalman filtering over single packet drop channel

Denote $\hat{x}(k|t)$ as the MMSE estimate of $x(k)$ and $\Sigma(k|t)$ as the corresponding error covariance at the data fusion center, conditioned on $\mathcal{S}(t) = \{s(\tau)\}_{\tau=0}^{t}$, for $k \geq t \geq 0$, respectively, i.e., $\hat{x}(k|t) = \mathrm{E}[x(k)|\mathcal{S}(t)]$, $\Sigma(k|t) = \mathrm{Cov}[x(k)|\mathcal{S}(t)]$. The following modified Kalman filtering algorithm for the dynamic system in (2.9) and (2.10) is proposed in [77]:

$$\begin{aligned}
L(k) &= \Sigma(k|k-1)C'[R + C\Sigma(k|k-1)C']^{-1}, \\
\hat{x}(k|k) &= \hat{x}(k|k-1) + L(k)[s(k) - \gamma(k)C\hat{x}(k|k-1)], \\
\Sigma(k|k) &= \Sigma(k|k-1) - \gamma(k)L(k)C\Sigma(k|k-1), \\
\hat{x}(k+1|k) &= A\hat{x}(k|k), \\
\Sigma(k+1|k) &= A\Sigma(k|k)A' + Q.
\end{aligned}$$

Note that the above is derived based on the assumption that $\gamma(k)$ is known at time k, in spite of it being a random variable. It is also demonstrated in [77] that there exists an infimum for p, denoted by p_{inf} and called critical data arrival rate, below which the estimation error covariance becomes unbounded. Under the TCP-like protocol, the successful transmission of the measurement packet is acknowledged at the data fusion center. Specifically $\gamma(k)$ is known at the fusion center and can be used to estimate $x(k+1)$. Therefore we consider the data fusion Kalman filter as the optimal one-step predictor. Furthermore we are interested in the stationary Kalman filter that minimizes the average error variance over the Bernoulli process $\{\gamma(k)\}$ in the steady-state at the data fusion center. Let $\widehat{x}(k)$ be the linear MMSE (LMMSE) estimate of $x(k)$ at the fusion center, conditioned on $\mathcal{S}(k-1)$ and averaged over $\{\gamma(t)\}_{t=0}^{k-1}$. The stationary Kalman filter that minimizes the average estimation error variance at the data fusion center is given by [74, 77]

$$\widehat{x}(k+1) = A\widehat{x}(k) + K\left[s(k) - \gamma(k)C\widehat{x}(k)\right], \tag{2.11a}$$

$$K = A\Sigma C'\left(R + C\Sigma C'\right)^{-1}, \tag{2.11b}$$

where Σ is the stabilizing solution to MARE

$$\Sigma = A\Sigma A' + Q - pA\Sigma C' \left(R + C\Sigma C'\right)^{-1} C\Sigma A'. \tag{2.12}$$

2.1.3 Critical Data Arrival Rate

Stationary Kalman filter (2.11) requires the existence of the stabilizing solution to MARE (2.12). It is important to observe that, different from the case of standard algebraic Riccati equation (ARE), the detectability of (C, A) and the stabilizability of $(A, Q^{1/2})$ on the unit circle are not adequate for MARE (2.12) to admit the stabilizing solution. As pointed out in [77], data arrival rate p plays a crucial role.

Denote the state estimation error at the data fusion center by

$$e_x(k) = x(k) - \widehat{x}(k).$$

Taking the difference between (2.9a) and (2.11a) yields estimation error dynamics

$$e_x(k+1) = \left[A - \gamma(k)KC\right] e_x(k) + w(k) - \gamma(k)Kv(k). \tag{2.13}$$

The MS stabilizability is defined next.

Definition 2. *Error dynamics* (2.13) *are said to be MS stabilizable, if for a given data arrival rate p, there exists K such that* (2.13) *is MS stable, i.e.,*

$$\lim_{k\to\infty} \mathrm{E}\left\{\|e_x(k)\|^2\right\} < \infty.$$

Such K is called an MS stabilizing gain.

Note that the two noise terms in (2.13) do not affect its MS stabilizability, and can be

removed for stability analysis. We thus consider the error dynamics described by

$$e_x(k+1) = [A - \gamma(k)KC]\, e_x(k). \tag{2.14}$$

In light of [74, 103], MARE (2.12) admits a unique stabilizing solution if and only if error dynamics (2.14) are MS stabilizable. Furthermore, MS stabilizability holds for (2.14) when $p > p_{\inf}$, while it fails when $p < p_{\inf}$. Hence we can find the critical data arrival rate $p_{\inf}$ by investigating the MS stabilizability condition for the error dynamics. Specifically we adopt

$$\gamma(k) = p\,[1 + \delta(k)]\,, \tag{2.15}$$

where $\{\delta(k)\}$ is a white random process with mean zero and variance

$$\mathrm{E}\{\delta(k)^2\} = \nu^2 = \frac{p - p^2}{p^2} = p^{-1} - 1 \iff p = \frac{1}{1+\nu^2}.$$

Substituting (2.15) into (2.14) yields the error dynamics in feedback form as

$$e_x(k+1) = (A + K_pC)e_x(k) + K_p[\delta(k)I_m]e_y(k), \;\; e_y(k) = Ce_x(k), \;\; K_p = -pK. \tag{2.16}$$

Feedback system (2.16) can be schematically illustrated by the block diagram in Figure 2.3. Now it is easy to see that the superium for variance ν^2, denoted by $\nu^2_{\sup}$, below which the MS stabilizability holds for error dynamics (2.14), is related to $p_{\inf}$ via $\nu^2_{\sup} = p_{\inf}^{-1} - 1$. It has been recognized [22, 93] that packet drop can be considered as a special case of multiplicative noise (fading channel). Hence the MS stability results introduced in Subsection 2.1.1 can be applied to study the stability issue of the error dynamics. The next result shows that the MS stabilizability for (2.14) is governed by some LMI.

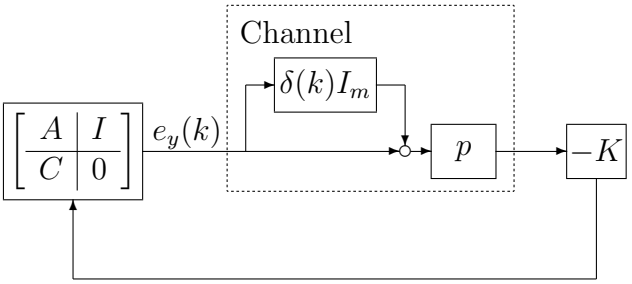


Figure 2.3: Estimation error dynamics in feedback form over single packet drop channel

Lemma 4. *Assume that (C, A) is detectable. Then for a given ν^2, error dynamics* (2.14) *are MS stabilizable, if and only if there exist $Z > 0$ and $K_{pZ} \in \mathbb{R}^{n\times m}$ satisfying the LMI:*

$$\begin{bmatrix} Z & ZA + K_{pZ}C & K_{pZ}C \\ A'Z + C'K'_{pZ} & Z & 0 \\ C'K'_{pZ} & 0 & \nu^{-2}Z \end{bmatrix} > 0.$$

If there exists a feasible solution pair (Z, K_{pZ}) to the above LMI, then $K_p = Z^{-1}K_{pZ}$ achieves MS stability for error dynamics (2.14).

Proof. The feedback system in Figure 2.3 can be converted to that in Figure 2.1 by taking

$$G(z) = C(zI - A - K_pC)^{-1}K_p.$$

By Lemma 1, the MS stabilizability is equivalent to the existence of $X_g > 0$ and $K_p \in \mathbb{R}^{n\times m}$ such that

$$X_g - (A + K_pC)X_g(A + K_pC)' - \nu^2 K_pCX_gC'K'_p > 0.$$

Multiplying $Z = X_g^{-1}$ from both left and right to the above inequality yields

$$Z - (ZA + K_{pZ}C)Z^{-1}(ZA + K_{pZ}C)' - \nu^2 K_{pZ}CZ^{-1}C'K'_{pZ} > 0,$$

where $K_{pZ} = ZK_p$. By the well-known *Schur Complement Lemma* [98], the above inequality and $Z > 0$ are in turn equivalent to the LMI in the lemma, thereby concluding the proof. □

Notice that Lemma 4 provides a way to verify if MS stabilizing gain exists for a given variance ν^2. However, in order to compute the critical data arrival rate $p_{\inf}$, we have to search for the maximum ν^2 over the feasible set of (Z, K_{pZ}) subject to the LMI, using either bisection or other algorithms iteratively. Upper and lower bounds of $p_{\inf}$ from [47, 77] are useful in such iterative searches. The following bounds

$$M(A)^{\frac{2}{m}} - 1 < \nu^{-2} < M(A)^2 - 1$$

can also be adopted, where $M(A) := \prod_{i=1}^{n} \max\{|\lambda_i(A)|, 1\}$ is the Mahler measure of matrix A. The next result presents a much more efficient algorithm to compute $\nu^2_{\sup}$ and $p_{\inf}$.

Theorem 1. *Under Assumption* 3 *and detectability of* (C, A)*, the superium* $\nu^2_{\sup}$ *can be computed via the minimization of* ν^{-2} *subject to the following LMIs over matrix pair* (X_Φ^{-1}, Φ^{-2})*:*

$$\begin{aligned}
&\text{(a) } \nu^{-2} X_\Phi^{-1} > C'\Phi^{-2}C, \\
&\text{(b) } X_\Phi^{-1} + C'\Phi^{-2}C \geq A' X_\Phi^{-1} A, \\
&\text{(c) } X_\Phi^{-1} > 0, \quad \Phi^{-2} > 0,
\end{aligned}$$

and the critical data arrival rate $p_{\inf} = (1 + \nu^2_{\sup})^{-1}$*.*

Proof. Recall that $G(z) = C(zI - A - K_pC)^{-1}K_p$. Denote $C_\Phi = \Phi^{-1}C$, $K_\Phi = K_p\Phi$, $A_K = A + K_\Phi C_\Phi$, and $G_\Phi(z) = C_\Phi(zI - A_K)^{-1}K_\Phi$. In light of Lemma 3, the MS stabilizability of the closed-loop system in Figure 2.3 is equivalent to the existence of $\Phi > 0$ such that

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_\Phi(e^{j\omega})G_\Phi(e^{j\omega})^* d\omega\right) < \nu^{-2}.$$

Let $Z_\Phi \geq 0$ be the solution to the Lyapunov equation

$$Z_\Phi = A_K Z_\Phi A_K' + K_\Phi K_\Phi'.$$

Note that decomposition can be applied to $G_\Phi(e^{j\omega})G_\Phi(e^{j\omega})^*$ to obtain

$$G_\Phi(e^{j\omega})G_\Phi(e^{j\omega})^* = C_\Phi Z_\Phi C_\Phi' + C_\Phi \left(e^{j\omega}I - A_K\right)^{-1} A_K Z_\Phi C_\Phi' + C_\Phi Z_\Phi A_K' \left(e^{-j\omega}I - A_K'\right)^{-1} C_\Phi'.$$

It follows that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} G_\Phi(e^{j\omega})G_\Phi(e^{j\omega})^* d\omega = C_\Phi Z_\Phi C_\Phi' \geq C_\Phi X_\Phi C_\Phi', \tag{2.17}$$

where $X_\Phi \geq 0$ is the stabilizing solution to ARE

$$X_\Phi = AX_\Phi(I + C_\Phi' C_\Phi X_\Phi)^{-1}A'.$$

Equality holds for (2.17), if and only if

$$K_\Phi = -AX_\Phi C_\Phi'(I + C_\Phi X_\Phi C_\Phi')^{-1} \iff K_p = K_\Phi\Phi^{-1} = -AX_\Phi C'(\Phi^2 + CX_\Phi C')^{-1}.$$

Denote $\mathcal{S}_K$ as the set of all K_Φ such that A_K is a Schur stability matrix. We thus have the equivalence of the MS stabilizability to

$$\inf_{K_\Phi\in\mathcal{S}_K,\Phi>0} \rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_\Phi(e^{j\omega})G_\Phi(e^{j\omega})^* d\omega\right) = \inf_{\Phi>0,X_\Phi>0} \rho\left(C_\Phi X_\Phi C_\Phi'\right) < \nu^{-2}.$$

The above is in turn equivalent to the existence of $X_\Phi > 0$ and $\Phi > 0$ subject to

$$C_\Phi X_\Phi C_\Phi' < \nu^{-2}I \iff \nu^{-2}\Phi^2 > CX_\Phi C' \tag{2.18}$$

and the ARE inequality

$$X_\Phi \geq A(X_\Phi^{-1} + C'\Phi^{-2}C)^{-1}A'. \tag{2.19}$$

Although (2.18) and (2.19) are not LMIs in terms of $X_\Phi > 0$ and $\Phi > 0$, they can be conver-

ted equivalently to LMIs over $X_\Phi^{-1} > 0$ and $\Phi^{-2} > 0$. Indeed (2.18) and (2.19) are equivalent to (a) and (b) respectively, thereby concluding the proof. □

Remark 1. The LMI problem in Theorem 1 belongs to the category of generalized eigenvalue minimization [13], aimed at the minimization of ν^{-2}, subject to linear fractional constraint (a), and general positivity constraints (b) and (c). Well-defined algorithms for the generalized eigenvalue minimization can thus be used to search for such a minimum value without lengthy and tedious iterations. In addition, the search for the MS stabilizing gain is no longer needed. However the algorithm presented in Theorem 1 does not yield the real superium ν_{sup}^2 due to finite digits used in any digital computers. That is, the maximum of ν^2, denoted by ν_{max}^2 and computed via the LMIs, is very close, but not equal to ν_{sup}^2. There holds $\nu_{\text{max}}^2 < \nu_{\text{sup}}^2$. As a result, the minimum of p, denoted by p_{min} and obtained via Theorem 1, is very close, but not equal to the exact critical arrival rate p_{inf}. There holds $p_{\text{min}} > p_{\text{inf}}$. □

Once p_{min} is obtained, the stabilizing solution to MARE (2.12) can be computed, provided that the MS stabilizability condition, i.e., $p \geq p_{\text{min}}$, holds, and subsequently the optimal MS stabilizing gain can be obtained following (2.11b). Given the MS stabilizability of error dynamics (2.14) and the detectability of (C, A), a sufficient condition for the existence of the stabilizing solution to MARE (2.12) is the stabilizability of $(A, Q^{1/2})$. Under both detectability of (C, A) and stabilizability of $(A, Q^{1/2})$, there exists a unique positive semidefinite solution Σ to MARE (2.12), and iterative algorithms can be used to compute Σ. Set the initial value $\widehat{\Sigma}(0) = \widehat{\Sigma}_0 \geq 0$, and compute $\widehat{\Sigma}(k)$ iteratively using the following modified difference Riccati equation (MDRE) [77]:

$$\widehat{\Sigma}(k+1) = A\widehat{\Sigma}(k)A' + Q - pA\widehat{\Sigma}(k)C'[R + C\widehat{\Sigma}(k)C']^{-1}C\widehat{\Sigma}(k)A'. \tag{2.20}$$

Then $\widehat{\Sigma}(k)$ approaches the stabilizing solution as $k \to \infty$, i.e.,

$$\Sigma = \lim_{k\to\infty} \widehat{\Sigma}(k).$$

It becomes more complicated, if $(A, Q^{1/2})$ is not stabilizable. In fact even for standard ARE, if

$$\operatorname{rank}\left\{\begin{bmatrix} zI - A & Q^{1/2} \end{bmatrix}\right\} = n, \ \ \forall\, |z| = 1,$$

then in general there exist more than one positive semi-definite solutions to the standard ARE, but only one of them is the stabilizing solution. A sufficiently large initial value has to be selected for the iterative process [30]. The iterative algorithm to compute the stabilizing solution to MARE (2.12) will be more difficult to analyze. An easy fix is to consider solving an approximate stabilizing solution to MARE (2.12) via computing iteratively the MDRE (2.20) by replacing Q with $Q_\phi = Q + \phi I$ for sufficiently small $\phi > 0$. This way ensures the controllability of $(A, Q_\phi^{1/2})$ and a unique positive definite solution in the limit.

2.2 Multiple Data Packet Drop Channels

In this section we consider the setting where distributed sensor nodes transmit their observations to the fusion center through multiple data packet drop channels. This scenario can be regarded as a generalization of the single packet drop channel case discussed in the previous section.

2.2.1 Mean Square Stability

Again we start by presenting the preliminaries on the MS stability condition for feedback systems with multiplicative noise over multiple feedback channels, which will be applied to derive the main results of this section. Consider the feedback system configured in Figure 2.4 with the plant model described by transfer matrix $G(z) = C_g(zI - A_g)^{-1}B_g$, where $A_g \in \mathbb{R}^{n\times n}$, $B_g \in \mathbb{R}^{n\times m}$, $C_g \in \mathbb{R}^{m\times n}$, specified by

$$B_g = \begin{bmatrix} B_{g;1} & \cdots & B_{g;N} \end{bmatrix}, \ \ B_{g;i} \in \mathbb{R}^{n\times m_i}, \ \ C_g = \begin{bmatrix} C_{g;1} \\ \vdots \\ C_{g;N} \end{bmatrix}, \ \ C_{g;i} \in \mathbb{R}^{m_i\times n}, \ \ \sum_{i=1}^{N} m_i = m.$$

Let $\mathcal{N} := \{1, \ldots, N\}$. The next assumption may be required in various situations.

Assumption 4. $\text{rank}\{B_{g;i}\} = \text{rank}\{C_{g;i}\} = m_i \ \forall\, i \in \mathcal{N}$.

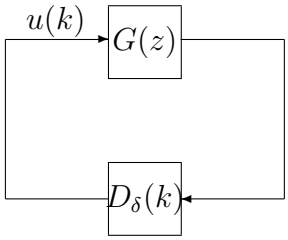


Figure 2.4: Feedback system with multiplicative noise over multiple feedback channels

Let the multiplicative noise (fading channels) in Figure 2.4 be specified by

$$D_\delta(k) = \text{diag}\left\{\delta_1(k) I_{m_1}, \ldots, \delta_N(k) I_{m_N}\right\},$$

where $\{\delta_i(k)\}_{i=1}^N$ are mutually independent white random processes with mean zero and variances $\{\sigma_i^2\}_{i=1}^N$, i.e.,

$$\text{E}\{\delta_i(k)\} = 0, \ \ \text{E}\{\delta_i(k)^2\} = \sigma_i^2 \ \ \forall\, i \in \mathcal{N}, \ \ k \geq 0.$$

The MS stability of the feedback system in Figure 2.4 is defined in Definition 1. The following MS stability condition is introduced in [13] as a generalization of that in Lemma 1.

Lemma 5. *Under Assumption* 1*, the following statements are equivalent:*

(i) *The feedback system in Figure* 2.4 *is MS stable;*

(ii) *There exists* $X_g > 0$ *satisfying the LMI:*

$$X_g > A_g X_g A_g' + \sum_{i=1}^N \sigma_i^2 B_{g;i} C_{g;i} X_g C_{g;i}' B_{g;i}'; \tag{2.21}$$

(iii) *There exists* $Y_g > 0$ *satisfying the LMI:*

$$Y_g > A_g' Y_g A_g + \sum_{i=1}^N \sigma_i^2 C_{g;i}' B_{g;i}' Y_g B_{g;i} C_{g;i}.$$

Building on Lemma 5, we have the following MS stability result, which will be applied directly to the derivation of the main results.

Lemma 6. *Under Assumptions* 1 *and* 4*, the following statements are equivalent:*

(i) *The feedback system in Figure* 2.4 *is MS stable;*

(ii) *There exists* $\Phi = \mathrm{diag}\,(\Phi_1, \ldots, \Phi_N) > 0$ *with* $\Phi_i \in \mathbb{R}^{m_i \times m_i}$ *and* $i \in \mathcal{N}$ *such that*

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_{\Phi;i}(e^{j\omega}) G_{\Phi;i}(e^{j\omega})^* d\omega\right) < \sigma_i^{-2}, \tag{2.22}$$

where $G_{\Phi;i}(z) = \Phi_i^{-1} C_{g;i}(zI - A_g)^{-1} B_g \Phi$*;*

(iii) *There exists* $\Psi = \mathrm{diag}\,(\Psi_1, \ldots, \Psi_N) > 0$ *with* $\Psi_i \in \mathbb{R}^{m_i \times m_i}$ *and* $i \in \mathcal{N}$ *such that*

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_{\Psi;i}(e^{j\omega})^* G_{\Psi;i}(e^{j\omega}) d\omega\right) < \sigma_i^{-2}, \tag{2.23}$$

where $G_{\Psi;i}(z) = \Psi C_g(zI - A_g)^{-1} B_{g;i} \Psi_i^{-1}$.

Proof. Denote $D_\sigma = \mathrm{diag}\,(\sigma_1 I_{m_1}, \ldots, \sigma_N I_{m_N})$.

For (i) $\Rightarrow$ (ii): There exists $X_g > 0$ satisfying LMI (2.21), if the feedback system in Figure 2.4 is MS stable, in light of Lemma 5. Substituting

$$X_g = z^{-1} A_g X_g + z X_g A_g' - A_g X_g A_g' + (zI - A_g) X_g (z^{-1} I - A_g')$$

into LMI (2.21) for $|z| = 1$ with rearrangement yields

$$\begin{aligned} M_g := \sum_{l=1}^{N} \sigma_l^2 B_{g;l} C_{g;l} X_g C_{g;l}' B_{g;l}' &< (zI - A_g) X_g (z^{-1} I - A_g') - 2 A_g X_g A_g' + z X_g A_g' + z^{-1} A_g X_g \\ &= (zI - A_g) X_g (z^{-1} I - A_g') + (zI - A_g) X_g A_g' + A_g X_g (z^{-1} I - A_g'). \end{aligned}$$

Denote $G_{i,l}(z) = C_{g;i}(zI - A_g)^{-1} B_{g;l}$. Then multiplying $C_{g;i}(zI - A_g)^{-1}$ from left and $(z^{-1} I -$

$A_g')^{-1}C_{g;i}'$ from right to the above inequality lead to

$$T_g(z) := C_{g;i}(zI - A_g)^{-1}M_g(z^{-1}I - A_g')^{-1}C_{g;i}' = \sum_{l=1}^{N} \sigma_l^2 G_{i,l}(z)C_{g;l}X_g C_{g;l}' G_{i,l}(z)^*$$
$$< C_{g;i}X_g C_{g;i}' + C_{g;i}X_g A_g'(z^{-1}I - A_g')^{-1}C_{g;i}' + C_{g;i}(zI - A_g)^{-1}A_g X_g C_{g;i}'.$$

Computing average over the unit circle gives

$$J := \frac{1}{2\pi}\int_{-\pi}^{\pi} T_g(e^{j\omega})\, d\omega = \sum_{l=1}^{N} \frac{\sigma_l^2}{2\pi}\int_{-\pi}^{\pi} G_{i,l}(e^{j\omega})C_{g;l}X_g C_{g;l}' G_{i,l}(e^{j\omega})^* d\omega < C_{g;i}X_g C_{g;i}'.$$

Denote $\Omega_l^2 = C_{g;l}X_g C_{g;l}'$ and $\Omega = \text{diag}\,(\Omega_1, \ldots, \Omega_N)$. Then $\Omega_l > 0$ and $\Omega > 0$ by $\text{rank}\{C_{g;l}\} = m_l$ and $X_g > 0$. The above inequality is equivalent to

$$I_{m_i} > \sum_{l=1}^{N} \frac{\sigma_l^2}{2\pi}\int_{-\pi}^{\pi} \Omega_i^{-1}G_{i,l}(e^{j\omega})\Omega_l^2 G_{i,l}(e^{j\omega})^*\Omega_i^{-1} d\omega = \frac{\sigma_i^2}{2\pi}\int_{-\pi}^{\pi} G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^* d\omega,$$

where $G_{\Phi;i}(z) = \Phi_i^{-1}C_{g;i}(zI - A_g)^{-1}B_g\Phi$ with $\Phi_i = \sigma_i\Omega_i$ and $\Phi = D_\sigma\Omega = \text{diag}\,(\Phi_1, \ldots, \Phi_N) > 0$. It follows that (2.22) holds $\forall\, i \in \mathcal{N}$.

For (ii) $\Rightarrow$ (i): By the Schur stability of A_g, there exists $X_g \geq 0$ to the Lyapunov equation

$$X_g = A_g X_g A_g' + \sum_{l=1}^{N} B_{g;l}\Phi_l^2 B_{g;l}'. \tag{2.24}$$

Similar to the derivation for $T_g(z)$, the power spectral density $\Psi_g(\omega) := G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^*$ admits the decomposition

$$\Psi_g(\omega) = \Phi_i^{-1}C_{g;i}X_g C_{g;i}'\Phi_i^{-1}$$
$$+ \Phi_i^{-1}C_{g;i}(e^{j\omega}I - A_g)^{-1}A_g X_g C_{g;i}'\Phi_i^{-1} + \Phi_i^{-1}C_{g;i}X_g A_g'(e^{-j\omega}I - A_g')^{-1}C_{g;i}'\Phi_i^{-1}.$$

If (2.22) holds, then X_g satisfies

$$\sigma_i^2 \Phi_i^{-1} C_{g;i} X_g C_{g;i}' \Phi_i^{-1} < I_{m_i}, \ \ \forall\, i \in \mathcal{N}. \tag{2.25}$$

Substituting (2.25) with $i = l$ into (2.24) yields

$$\begin{aligned} X_g &> A_g X_g A_g' + \sum_{l=1}^{N} \sigma_l^2 B_{g;l} \Phi_l (\Phi_l^{-1} C_{g;l} X_g C_{g;l}' \Phi_l^{-1}) \Phi_l B_{g;l}' \\ &= A_g X_g A_g' + \sum_{l=1}^{N} \sigma_l^2 B_{g;l} C_{g;l} X_g C_{g;l}' B_{g;l}', \end{aligned}$$

which further implies $X_g > 0$, thereby concluding the proof for (2.22). Since (2.23) is dual to (2.22), its proof is omitted. □

2.2.2 Stationary Data Fusion Kalman Filter

We now study the data fusion Kalman filtering in the presence of packet drops for networked system described by

$$x(k+1) = Ax(k) + w(k), \tag{2.26a}$$

$$y_i(k) = C_i x(k) + v_i(k), \ \ i \in \mathcal{N}, \tag{2.26b}$$

where $x(k) \in \mathbb{R}^n$ denotes the state of the dynamic target system, and $y_i(k) \in \mathbb{R}^{m_i}$ denotes the measurement obtained by the ith sensor node at time index k with $C_i \in \mathbb{R}^{m_i \times n}$ different for each $i \in \mathcal{N}$ generally. The process noise $w(k)$ and the measurement noise $v_i(k)$ are independent white processes with mean zero and covariance $Q \geq 0$ and $R_i > 0$, respectively, uncorrelated to the initial state $x(0) = x_0$ that has mean $\overline{x}_0$ and covariance Σ_0. By convention we assume that x_0, $w(k)$, $\{v_i(k)\}_{i=1}^N$ are jointly Gaussian distributed.

Due to the nature of WSNs, the signal received at the data fusion center from the ith sensor node is given by

$$s_i(k) = \gamma_i(k) y_i(k), \quad i \in \mathcal{N}, \tag{2.27}$$

where $\gamma_i(k)$ is imposed to represent the network distortion induced by data packet drops. Again we adopt a simple model of i.i.d. stationary Bernoulli process for $\{\gamma_i(k)\}$ with probability

$$\mathrm{P}\{\gamma_i(k) = 1\} = p_i > 0 \ \ \forall\, k \geq 0.$$

It follows that

$$\mathrm{E}\{\gamma_i(k)\} = p_i, \quad \mu_i^2 = \mathrm{E}\{[\gamma_i(k) - p_i]^2\} = p_i - p_i^2.$$

It is also assumed that $\gamma_i(k)$ is independent of the noise and the initial state. Based on (2.27), measurement $y_i(k)$ reaches the data fusion center at time k if $\gamma_i(k) = 1$; otherwise all the components of $y_i(k)$ are lost. The complete setup is depicted in Figure 2.5. Denote

$$y(k) = \mathrm{vec}\{y_1(k), \ldots, y_N(k)\}, \quad v(k) = \mathrm{vec}\{v_1(k), \ldots, v_N(k)\},$$
$$s(k) = \mathrm{vec}\{s_1(k), \ldots, s_N(k)\}, \quad C = \mathrm{col}\{C_1, \ldots, C_N\},$$
$$R = \mathrm{E}\{v(k)v(k)'\} = \mathrm{diag}(R_1, \ldots, R_N), \quad D_\gamma(k) = \mathrm{diag}\left\{\gamma_1(k) I_{m_1}, \ldots, \gamma_N(k) I_{m_N}\right\}.$$

Then the state space model in (2.26) and (2.27) can be written in a compact form as

$$x(k+1) = Ax(k) + w(k), \tag{2.28a}$$
$$y(k) = Cx(k) + v(k), \tag{2.28b}$$
$$s(k) = D_\gamma(k) y(k). \tag{2.28c}$$

Assumption 5. $\mathrm{rank}\{C_i\} = m_i \ \ \forall\, i \in \mathcal{N}$.

The above assumption is made without loss of generality. If $\mathrm{rank}\{C_i\} = \tilde{m}_i < m_i$, then measurement of $C_i x(k)$ at the ith sensor node has redundancies that can be removed in

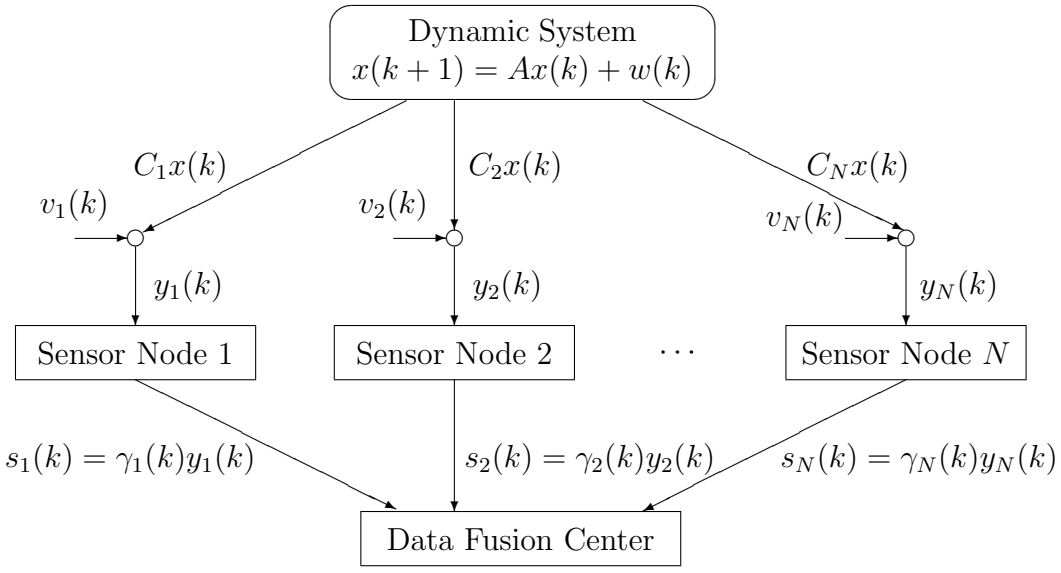


Figure 2.5: Setup for data fusion Kalman filtering over multiple packet drop channels

order to reduce the sensing and transmitting cost. Specifically in this case, there exists an orthogonal matrix $\Theta_i \in \mathbb{R}^{m_i \times \tilde{m}_i}$, satisfying $\Theta_i'\Theta_i = I_{\tilde{m}_i}$ and $C_i = \Theta_i\tilde{C}_i$, where $\text{rank}\{\tilde{C}_i\} = \tilde{m}_i$. We can thus transmit measurement $\tilde{y}_i(k) = \Theta_i'y_i(k) \in \mathbb{R}^{\tilde{m}_i}$ that has lower dimension.

Denote $\mathcal{S}(k) = \{s(t)\}_{t=0}^k$, $\mathcal{D}_\gamma(k) = \{D_\gamma(t)\}_{t=0}^k$. Let $\hat{x}(k)$ be the MMSE estimate of $x(k)$ and $\Sigma(k)$ be the corresponding error covariance at the data fusion center, conditioned on $\mathcal{S}(k-1)$, respectively, i.e., $\hat{x}(k) = \text{E}[x(k)|\mathcal{S}(k-1)]$, $\Sigma(k) = \text{Cov}[x(k)|\mathcal{S}(k-1)]$. Under the TCP-like protocol, the time-varying data fusion Kalman filter has the structure of an optimal one-step predictor as

$$\hat{x}(k+1) = A\hat{x}(k) + K(k)\left[s(k) - D_\gamma(k)C\hat{x}(k)\right]. \tag{2.29}$$

We are interested in deriving the stationary Kalman filter that minimizes the average error variance over the Bernoulli processes $\{\gamma_i(k)\}_{i=1}^N$ in the steady-state at the data fusion center. Let $\widehat{x}(k)$ be the LMMSE estimate of $x(k)$ at the fusion center, conditioned on $\mathcal{S}(k-1)$ and averaged over $\mathcal{D}_\gamma(k-1)$, and $\widehat{\Sigma}(k)$ be the average error covariance associated with $\widehat{x}(k)$. The stationary data fusion Kalman filter is presented in the following theorem.

Theorem 2. *For the system dynamics described in* (2.28)*, assume that there exists a stabilizing solution* Σ *to MARE*

$$\Sigma = A\Sigma A' + Q - A\Sigma C' \left[W \circ (R + C\Sigma C')\right]^{-1} C\Sigma A', \tag{2.30}$$

where $W = \mathbf{1}_m \mathbf{1}_m' + D_p^{-1} D_{\mu^2} D_p^{-1}$ *with*

$$D_p = \mathrm{E}\left\{D_\gamma(k)\right\} = \mathrm{diag}(p_1 I_{m_1}, \ldots, p_N I_{m_N}), \quad D_{\mu^2} = \mathrm{diag}(\mu_1^2 \mathbf{1}_{m_1} \mathbf{1}_{m_1}', \ldots, \mu_N^2 \mathbf{1}_{m_N} \mathbf{1}_{m_N}').$$

The stationary Kalman filter that minimizes the average error variance at the data fusion center is specified by

$$\widehat{x}(k+1) = A\widehat{x}(k) + K\left[s(k) - D_\gamma(k) C\widehat{x}(k)\right], \tag{2.31a}$$

$$K = A\Sigma C' \left[W \circ (R + C\Sigma C')\right]^{-1} D_p^{-1}. \tag{2.31b}$$

Proof. Following the conventional method, the optimal time-varying estimation gain $K(k)$ is used in deriving MDRE prior to taking $k \to \infty$ to obtain the optimal stationary estimation gain K and the associated MARE. Denote $e(k) = x(k) - \hat{x}(k)$ as the state estimation error. Taking the difference between (2.28a) and (2.29) yields error dynamics

$$e(k+1) = \left[A - K(k) D_\gamma(k) C\right] e(k) + w(k) - K(k) D_\gamma(k) v(k).$$

Evaluation of the average covariance of $e(k)$ can be carried out in two steps. First, taking covariance of $e(k)$ by treating $D_\gamma(k)$ as a deterministic time-varying matrix quantity yields

$$\Sigma(k+1) = \left[A - K(k) D_\gamma(k) C\right] \Sigma(k) \left[A - K(k) D_\gamma(k) C\right]' + Q + K(k) D_\gamma(k) R D_\gamma(k) K(k)'$$

in light of the independence of $w(k)$, $v(k)$, and x_0. The above can be rewritten as

$$\Sigma(k+1) = A\Sigma(k)A' - K(k)D_\gamma(k)C\Sigma(k)A' - A\Sigma(k)C'D_\gamma(k)K(k)' + Q$$
$$+ K(k)D_\gamma(k)\left[R + C\Sigma(k)C'\right]D_\gamma(k)K(k)'.$$

Denote $\mathrm{E}_\gamma\{\cdot\}$ as the expectation with respect to $D_\gamma(k)$. In the second step, the $\mathrm{E}_\gamma\{\cdot\}$ operation can be taken to the above equation, leading to

$$\widehat{\Sigma}(k+1) = A\widehat{\Sigma}(k)A' - K(k)D_pC\widehat{\Sigma}(k)A' - A\widehat{\Sigma}(k)C'D_pK(k)' + Q$$
$$+ K(k)\mathrm{E}_\gamma\left\{D_\gamma(k)\left[R + C\widehat{\Sigma}(k)C'\right]D_\gamma(k)\right\}K(k)'. \tag{2.32}$$

Notice that $\widehat{\Sigma}(k) = \mathrm{E}_\gamma\{\Sigma(k)\}$. By direct calculations,

$$\begin{aligned}
&\mathrm{E}_\gamma\left\{D_\gamma(k)\left[R + C\widehat{\Sigma}(k)C'\right]D_\gamma(k)\right\} \\
&= D_p\left[R + C\widehat{\Sigma}(k)C'\right]D_p + \mathrm{E}_\gamma\left\{[D_\gamma(k) - D_p]\left[R + C\widehat{\Sigma}(k)C'\right][D_\gamma(k) - D_p]\right\} \\
&= D_p\left[R + C\widehat{\Sigma}(k)C'\right]D_p + D_{\mu^2} \circ \left[R + C\widehat{\Sigma}(k)C'\right] = D_p\left\{W \circ \left[R + C\widehat{\Sigma}(k)C'\right]\right\}D_p.
\end{aligned} \tag{2.33}$$

Substituting (2.33) into (2.32) yields

$$\widehat{\Sigma}(k+1) = A\widehat{\Sigma}(k)A' + Q - A\widehat{\Sigma}(k)C'\left\{W \circ \left[R + C\widehat{\Sigma}(k)C'\right]\right\}^{-1}C\widehat{\Sigma}(k)A'$$
$$+ \Pi(k)\left\{W \circ \left[R + C\widehat{\Sigma}(k)C'\right]\right\}^{-1}\Pi(k)', \tag{2.34}$$

where

$$\Pi(k) = A\widehat{\Sigma}(k)C' - K(k)D_p\left\{W \circ \left[R + C\widehat{\Sigma}(k)C'\right]\right\}.$$

Since the last term in (2.34) is nonnegative definite and is the only term involving the

estimation gain, setting $\Pi(k) = 0$ yields the optimal time-varying estimation gain

$$K(k) = A\widehat{\Sigma}(k)C' \left\{ W \circ \left[R + C\widehat{\Sigma}(k)C' \right] \right\}^{-1} D_p^{-1}.$$

Equation (2.34) is thus reduced to MDRE

$$\widehat{\Sigma}(k+1) = A\widehat{\Sigma}(k)A' + Q - A\widehat{\Sigma}(k)C' \left\{ W \circ \left[R + C\widehat{\Sigma}(k)C' \right] \right\}^{-1} C\widehat{\Sigma}(k)A', \tag{2.35}$$

which is the minimized error covariance. As $k \to \infty$, MDRE (2.35) converges to MARE (2.30), and $K(k)$ converges to the estimation gain K in (2.31b). It follows that (2.31) is indeed the stationary Kalman filter that minimizes the average error variance. □

We claim that the stationary Kalman filter (2.31) is consistent with the existing results on Kalman filtering with intermittent observations. When $N = 1$, the multiple sensor nodes (packet drop channels) setting degrades to the single sensor node (packet drop channel) case studied in Section 2.1, where only a single random process $\{\gamma(k)\}$, instead of $\{\gamma_i(k)\}_{i=1}^N$, is adopted. In that case, $D_p = pI_m$, and $W = (1 + p^{-2}\mu^2)\mathbf{1}_m\mathbf{1}_m' = p^{-1}\mathbf{1}_m\mathbf{1}_m'$. Then MARE (2.30) reduces to MARE (2.12), and the optimal estimation gain in (2.31b) is equivalent to that in (2.11b), which agree with the results in [77]. When $m_i = 1$ for all $i \in \mathcal{N}$, $y_i(k)$, the data packet measured and transmitted by sensor node i, has dimension one for all $i \in \mathcal{N}$. Then our networked system reduces to the dual of the system studied in [102]. In that case, $D_p = \mathrm{diag}(p_1, \ldots, p_N)$,

$$W = \begin{bmatrix} 1 + \mathrm{SNR}_1^{-1} & 1 & \cdots & 1 \\ 1 & 1 + \mathrm{SNR}_2^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1 + \mathrm{SNR}_N^{-1} \end{bmatrix},$$

where $\mathrm{SNR}_i = p_i^2/\mu_i^2$ is the signal-to-noise ratio of the ith packet drop channel. Hence the resulting MARE (2.30) and optimal gain (2.31b) are dual to the ones derived in [102].

2.2.3 Stability Margin

The stationary Kalman filter (2.31) proposed in Theorem 2 assumes the existence of the stabilizing solution to MARE (2.30), which is hinged on the MS stabilizability of the estimation error dynamics, similar to the case of MARE (2.12). The estimation error dynamics associated to filter (2.31) are given by

$$e_x(k+1) = [A - KD_\gamma(k)C]\, e_x(k), \tag{2.36}$$

with noise terms removed. Specially we adopt

$$D_\gamma(k) = D_p\,[I + D_\delta(k)]\,, \quad D_\delta(k) = \mathrm{diag}\,\{\delta_1(k)I_{m_1}, \ldots, \delta_N(k)I_{m_N}\}\,, \tag{2.37}$$

where $\{\delta_i(k)\}_{i=1}^N$ are mutually independent white random processes with mean zero and variance

$$\mathrm{E}\{\delta_i(k)^2\} = \nu_i^2 = \frac{\mu_i^2}{p_i^2} = p_i^{-1} - 1 \iff p_i = \frac{1}{1+\nu_i^2}.$$

Substituting (2.37) into (2.36) yields the error dynamics in feedback form as

$$e_x(k+1) = (A + K_pC)e_x(k) + K_pD_\delta(k)e_y(k), \quad e_y(k) = Ce_x(k), \quad K_p = -KD_p. \tag{2.38}$$

Feedback system (2.38) can be schematically illustrated by the block diagram in Figure 2.6.

In light of [74, 103], MARE (2.30) admits a unique stabilizing solution if and only if MS stabilizability holds for error dynamics (2.36). For the single sensor node case discussed in the previous section, there exists a superium for ν^2, denoted by ν^2_{sup}, below which the corresponding error dynamics are MS stabilizable. However when multiple packet drop channels are involved, such superiums for individual ν_i^2 no longer exist. Alternatively, we

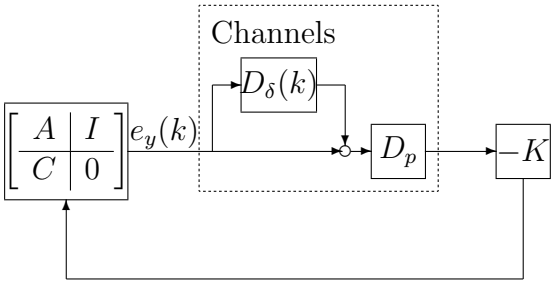


Figure 2.6: Estimation error dynamics in feedback form over multiple packet drop channels

adopt the stability margin, termed as $\alpha_{\max} > 0$, to help to determine the MS stabilizability. The next result presents an efficient algorithm to compute $\alpha_{\max}$.

Theorem 3. *Under Assumption* 5 *and the detectability of* (C, A)*, for a given set of variances* $\{\nu_i^2\}_{i=1}^N$*, the stability margin* $\alpha_{\max} = \beta_{\min}^{-1}$ *can be computed via the minimization of* β *subject to the following LMIs over matrix pair* (X_Φ^{-1}, Φ^{-2})*:*

$$\begin{aligned}
&\text{(a) } \beta\nu_i^{-2} X_\Phi^{-1} > C_i' \Phi_i^{-2} C_i, \ \ \forall\, i \in \mathcal{N}, \\
&\text{(b) } X_\Phi^{-1} + C'\Phi^{-2}C \geq A' X_\Phi^{-1} A, \\
&\text{(c) } X_\Phi^{-1} > 0, \quad \Phi^{-2} > 0,
\end{aligned}$$

where $\Phi = \text{diag}(\Phi_1, \ldots, \Phi_N)$*. Furthermore, error dynamics* (2.36) *are MS stabilizable, if and only if* $\alpha_{\max} = \beta_{\min}^{-1} \geq 1$*.*

Proof. The feedback system in Figure 2.6 can be converted to that in Figure 2.4 by taking $G(z) = C(zI - A - K_p C)^{-1} K_p$. Denote $C_{\Phi;i} = \Phi_i^{-1} C_i$, $C_\Phi = \Phi^{-1} C$, $K_\Phi = K_p \Phi$, $A_K = A + K_\Phi C_\Phi$, and $G_{\Phi;i}(z) = C_{\Phi;i}(zI - A_K)^{-1} K_\Phi$, $i \in \mathcal{N}$. In light of Lemma 6, the MS stabilizability of the closed-loop system in Figure 2.6 is equivalent to the existence of $\Phi = \text{diag}\,(\Phi_1, \ldots, \Phi_N) > 0$ such that

$$\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_{\Phi;i}(e^{j\omega}) G_{\Phi;i}(e^{j\omega})^* d\omega\right) < \nu_i^{-2} \ \ \forall\, i \in \mathcal{N}.$$

Let $Z_\Phi \geq 0$ be the solution to the Lyapunov equation

$$Z_\Phi = A_K Z_\Phi A_K' + K_\Phi K_\Phi'.$$

Note that decomposition can be applied to $G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^*$ to obtain

$$\begin{aligned} G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^* = C_{\Phi;i}Z_\Phi C_{\Phi;i}' + C_{\Phi;i}\left(e^{j\omega}I - A_K\right)^{-1} A_K Z_\Phi C_{\Phi;i}' \\ + C_{\Phi;i}Z_\Phi A_K'\left(e^{-j\omega}I - A_K'\right)^{-1} C_{\Phi;i}'. \end{aligned}$$

It follows that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^* d\omega = C_{\Phi;i}Z_\Phi C_{\Phi;i}' \geq C_{\Phi;i}X_\Phi C_{\Phi;i}', \tag{2.39}$$

where $X_\Phi \geq 0$ is the stabilizing solution to ARE

$$X_\Phi = AX_\Phi(I + C_\Phi' C_\Phi X_\Phi)^{-1}A'.$$

Equality holds for (2.39), if and only if

$$K_\Phi = -AX_\Phi C_\Phi'(I + C_\Phi X_\Phi C_\Phi')^{-1} \iff K_p = K_\Phi\Phi^{-1} = -AX_\Phi C'(\Phi^2 + CX_\Phi C')^{-1}.$$

Denote $\mathcal{S}_K$ as the set of all K_Φ such that A_K is a Schur stability matrix. We thus have the equivalence of the MS stabilizability to

$$\inf_{K_\Phi\in\mathcal{S}_K,\Phi>0}\rho\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} G_{\Phi;i}(e^{j\omega})G_{\Phi;i}(e^{j\omega})^* d\omega\right) = \inf_{\Phi>0,X_\Phi>0}\rho\left(C_{\Phi;i}X_\Phi C_{\Phi;i}'\right) < \nu_i^{-2} \ \ \forall\, i \in \mathcal{N}.$$

The above is in turn equivalent to the existence of $X_\Phi > 0$ and $\Phi > 0$ subject to

$$C_{\Phi;i}X_\Phi C_{\Phi;i}' < \nu_i^{-2} I_{m_i} \iff \nu_i^{-2}\Phi_i^2 > C_i X_\Phi C_i' \tag{2.40}$$

and the ARE inequality

$$X_\Phi \geq A(X_\Phi^{-1} + C'\Phi^{-2}C)^{-1}A'. \tag{2.41}$$

Although (2.40) and (2.41) are not LMIs in terms of $X_\Phi > 0$ and $\Phi > 0$, they can be converted equivalently to LMIs over $X_\Phi^{-1} > 0$ and $\Phi^{-2} > 0$. Indeed (2.40) is equivalent to

$$\nu_i^{-2} X_\Phi^{-1} > C_i' \Phi_i^{-2} C_i, \tag{2.42}$$

while (2.41) is equivalent to (b). Then the stability margin $\alpha_{\max} = \beta_{\min}^{-1}$ can be obtained by minimizing β subject to LMIs (a), (b), and (c) in the theorem. It is easy to see that there exists a feasible solution pair (X_Φ^{-1}, Φ^{-2}) to LMIs (2.42), (b), and (c), if and only if $\beta_{\min} \leq 1$, i.e., $\alpha_{\max} \geq 1$, which concludes the proof. □

Remark 2. Similar to Theorem 1, Theorem 3 also provides an LMI based solution, which belongs to the category of generalized eigenvalue minimization. Hence well-defined algorithms for the generalized eigenvalue minimization can be used for computation. We comment that $\alpha_{\max} < \alpha_{\sup}$. The gap between $\alpha_{\max}$ and $\alpha_{\sup}$ is determined by numerical precision. For the given variances $\{\nu_i^2\}_{i=1}^N$, if $\alpha_{\max} > 1$, then the MS stabilizability holds even if the variances are increased to $\{\alpha_{\max}\nu_i^2\}_{i=1}^N$, which explains why we call $\alpha_{\max}$ stability margin. □

In light of [103], given a set of variances $\{\nu_i^2\}_{i=1}^N$, the stabilizing solution to MARE (2.30) can be computed, provided that the MS stabilizability condition, i.e., $\alpha_{\max} \geq 1$, holds, and subsequently the optimal MS stabilizing gain can be obtained following (2.31b). The procedure to compute the stabilizing solution to MARE (2.30) is basically identical to that for MARE (2.12), thus is omitted here. Details can be referred to Subsection 2.1.3.

2.3 Numerical Examples

Two numerical examples are presented to illustrate how to apply Theorem 1 and Theorem 3 to address the MS stabilizability issue.

Example 2.1. The first example is adapted from the example in [74]. We consider the pendubot system: a classic control laboratory experiment. More details and references on the pendubot can be found in [101]. The linearized and discretized state space model is given by $x(k+1) = Ax(k) + Bu(k) + w(k)$ with

$$A = \begin{bmatrix} 1.001 & 0.005 & 0.000 & 0.000 \\ 0.350 & 1.001 & -0.135 & 0.000 \\ -0.001 & 0.000 & 1.001 & 0.005 \\ -0.375 & -0.001 & 0.590 & 1.001 \end{bmatrix}.$$

Notice that the impact of input term $Bu(k)$ on the state dynamics can always be offset by adding this term to the estimator. Thus the input term will not affect the MS stabilizability of the error dynamics, and can be disregarded for computation convenience. System matrix A has two unstable eigenvalues, i.e., 1.061, 1.033. A single sensor node is employed to observe the pendubot system and send its measurements to the data fusion center for state estimation and tracking. The observation of the node follows model (2.9b) with

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The noise statistics are given by

$$Q = qq', \quad q = \begin{bmatrix} 0.003 & 1.000 & -0.005 & -2.150 \end{bmatrix}', \quad R = 0.001 I_2.$$

It is easy to verify that $(A, Q^{1/2})$ is controllable and (C, A) is observable.

Since only a single sensor node is used in this example, we can apply Theorem 1 to find the critical data arrival rate for the communication channel from the sensor node to the estimation center. Solving the LMI problem in Theorem 1 yields $\nu_{\max}^{-2} = 0.1262$, and thus $p_{\min} = 0.1121$, which can be considered as the critical data arrival rate. Then we know that

MARE (2.12) admits a unique stabilizing solution when $p = 0.2 > p_{\min}$. Initialize MDRE (2.20) with $\widehat{\Sigma}(0) = 0$ and take the limit. The resulting stabilizing solution is given by

$$\Sigma = \begin{bmatrix} 0.0205 & 0.4507 & -0.1789 & -2.0475 \\ 0.4507 & 14.2728 & -6.6573 & -75.3820 \\ -0.1789 & -6.6573 & 7.4484 & 69.1820 \\ -2.0475 & -75.3820 & 69.1820 & 665.8969 \end{bmatrix}.$$

When $p = 0.1 < p_{\min}$, $\widehat{\Sigma}(k)$ becomes unbounded as $k \to \infty$, implying that error dynamics (2.14) cannot be made MS stable given any estimation gain if the actual data arrival rate of the communication channel is lower than the critical rate. It is also worth noting that when the arrival rate is higher than but close to the critical threshold, the convergence of $\widehat{\Sigma}(k)$ to Σ becomes very slow. In general, higher arrival rates lead to faster convergence.

Example 2.2. In the second example, we consider a system of dynamics (2.28a) with system matrix A specified by

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Obviously A has four unstable eigenvalues. A total of $N = 4$ sensor nodes in a WSN are employed to observe the dynamic system and send their measurements to the data fusion center for state estimation. The observation matrix for each sensor node is specified by

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \; C_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; C_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}, \; C_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

It is important to observe that each node can only observe a partial of the state, thus the

target is undetectable by individual nodes. However the estimation center receives measurement information from all the four nodes, which guarantees the detectability of (C, A), where $C = [C_1' \; C_2' \; C_3' \; C_4']'$. The noise statistics are given by

$$Q = 0.002 I_4, \;\; R_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \;\; R_2 = \sqrt{2}\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \;\; R_3 = \sqrt{3}, \;\; R_4 = \sqrt{4},$$

which implies the stabilizability of $(A, Q^{1/2})$.

Since multiple sensor nodes are used in this example, Theorem 3 can be applied to find the stability margin given a set of packet drop statistics, which helps us to tell if MS stabilizability holds for the corresponding error dynamics. Consider the case where

$$p_1 = 0.9, \quad p_2 = 0.7, \quad p_3 = 0.5, \quad p_4 = 0.3.$$

Then we have $\nu_1^{-2} = 9$, $\nu_2^{-2} = 2.3333$, $\nu_3^{-2} = 1$, $\nu_4^{-2} = 0.4286$. Solving the LMI problem in Theorem 3 yields $\alpha_{\max} = 3.0000 > 1$, which implies that error dynamics (2.36) are MS stabilizable, and MARE (2.30) admits a unique stabilizing solution. Initializing MDRE (2.35) with $\widehat{\Sigma}(0) = 0$ yields the limit that is the stabilizing solution to MARE, given by

$$\Sigma = \begin{bmatrix} 2.8181 & 1.4073 & 0.0234 & 0 \\ 1.4073 & 4.4578 & 0.0117 & 0 \\ 0.0234 & 0.0117 & 0.0493 & 0 \\ 0 & 0 & 0 & 0.1189 \end{bmatrix}.$$

If the data arrival rates for individual communication channels are changed to

$$p_1 = 0.7, \quad p_2 = 0.9, \quad p_3 = 0.5, \quad p_4 = 0.3,$$

then the resulting stability margin $\alpha_{\max} = 0.7778 < 1$, implying that error dynamics (2.36)

cannot be made MS stable under the given statistics. Although the only difference between these two sets of data arrival rates is that the values for p_1 and p_2 are switched, there is a big gap between the two resulting stability margins. A close look at the structure of matrix A shows that the second entry of the state is governed by the unstable eigenvalue 2, while it is observable only by sensor node 1. Therefore a sufficiently large data arrival rate p_1 should be assigned to the first channel so that frequent observations from node 1 can be supplied to the estimation center to track the fast-changing second state component. Otherwise the estimation error may become unbounded as time proceeds. It is also worth noting that when the stability margin is greater than but close to 1, the convergence of $\widehat{\Sigma}(k)$ to Σ becomes very slow. Generally a greater $\alpha_{\max}$ leads to faster convergence.

CHAPTER 3
KALMAN CONSENSUS FILTER OVER DATA PACKET DROP CHANNELS

In this chapter we investigate distributed Kalman filtering over WSNs, where no data fusion center exists and each sensor node is required to locally estimate the state of a discrete-time dynamic system based on its own measurements and the information transmitted from its neighbors in the presence of data packet drops. The stationary DKF is proposed first, followed by the design of the KCF.

3.1 Distributed Kalman Filter

Prior to tackling the KCF, we examine a general MMSE estimation problem, and apply the results to derive the stationary DKF, which can be regarded as a special case of the KCF without the consensus term.

3.1.1 Distributed Estimation

We consider static distributed estimation of vector $X \in \mathbb{R}^n$ using a wireless sensor network consisting of N nodes, indexed sequentially by the index set $\mathcal{N} := \{1, \ldots, N\}$. It is assumed that sensor node i not only has local observation $Y_i \in \mathbb{R}^{m_i}$, but also may receive observations $\{Y_j\}_{j\in\mathcal{N}_i}$, transmitted from its neighbors specified by index subset $\mathcal{N}_i \subset \mathcal{N}$. Denote $\mathcal{J}_i$ as the union of $\mathcal{N}_i$ and node i itself, i.e., $\mathcal{J}_i = \mathcal{N}_i \cup \{i\}$. It is assumed that, at node i, X and $Y_{\mathcal{J}_i} = \text{vec}\{Y_j\}_{j\in\mathcal{J}_i}$ are modeled as two jointly Gaussian distributed random vectors, having statistics

$$\text{E}\left\{\begin{bmatrix} X \\ Y_{\mathcal{J}_i} \end{bmatrix}\right\} = \begin{bmatrix} \overline{X}_i \\ \overline{Y}_{\mathcal{J}_i} \end{bmatrix}, \quad \text{Cov}\left\{\begin{bmatrix} X \\ Y_{\mathcal{J}_i} \end{bmatrix}\right\} = \begin{bmatrix} \Sigma_{x_i,x_i} & \Sigma_{x_i,y_{\mathcal{J}_i}} \\ \Sigma_{y_{\mathcal{J}_i},x_i} & \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}} \end{bmatrix}.$$

We aim at computing the MMSE estimate of X based on the collective observations $Y_{\mathcal{J}_i}$ at sensor node i. Due to the presence of data packet drops, the observed signal received at the

ith sensor node from its neighboring nodes in $\mathcal{N}_i$ are given by $\{S_{i,j} = \gamma_{i,j} Y_j\}_{j \in \mathcal{N}_i}$ instead of $\{Y_j\}_{j \in \mathcal{N}_i}$, where $\{\gamma_{i,j}\}_{j \in \mathcal{N}_i}$ represent the network distortion induced by data packet drops. Specifically $\gamma_{i,j} = 1$, if Y_j is received, and $\gamma_{i,j} = 0$, if Y_j is not received by the ith node for each $j \in \mathcal{N}_i$. Because of the use of the TCP-like protocol, the value of $\gamma_{i,j}$ is acknowledged at the ith node; Thus it is treated as a deterministic quantity in deriving the MMSE estimate of X. Since Y_i is directly measured by the ith node without the need for transmission, $\gamma_{i,i} = 1$ is assumed. Let $D_{\gamma;\mathcal{J}_i} = \text{diag}\{\gamma_{i,j} I_{m_j}\}_{j \in \mathcal{J}_i}$ with known $\{\gamma_{i,j}\}_{j \in \mathcal{N}_i}$. Then X and $S_{\mathcal{J}_i} = \text{vec}\{S_{i,j}\}_{j \in \mathcal{J}_i} = D_{\gamma;\mathcal{J}_i} Y_{\mathcal{J}_i}$ are two jointly Gaussian distributed random vectors, having statistics

$$\text{E}\left\{ \begin{bmatrix} X \\ S_{\mathcal{J}_i} \end{bmatrix} \right\} = \begin{bmatrix} \overline{X}_i \\ \overline{S}_{\mathcal{J}_i} \end{bmatrix} = \begin{bmatrix} \overline{X}_i \\ D_{\gamma;\mathcal{J}_i} \overline{Y}_{\mathcal{J}_i} \end{bmatrix},$$

$$\text{Cov}\left\{ \begin{bmatrix} X \\ S_{\mathcal{J}_i} \end{bmatrix} \right\} = \begin{bmatrix} \Sigma_{x_i,x_i} & \Sigma_{x_i,s_{\mathcal{J}_i}} \\ \Sigma_{s_{\mathcal{J}_i},x_i} & \Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} \end{bmatrix} = \begin{bmatrix} \Sigma_{x_i,x_i} & \Sigma_{x_i,y_{\mathcal{J}_i}} D_{\gamma;\mathcal{J}_i} \\ D_{\gamma;\mathcal{J}_i} \Sigma_{y_{\mathcal{J}_i},x_i} & D_{\gamma;\mathcal{J}_i} \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}} D_{\gamma;\mathcal{J}_i} \end{bmatrix}$$

at the ith node.

Denote $\hat{X}_i$ as the local MMSE estimator of X at node i, conditioned on $S_{\mathcal{J}_i}$. The Gaussian assumption implies that $\hat{X}_i = \text{E}\{X | S_{\mathcal{J}_i}\}$. Let $X_{e;i} = X - \hat{X}_i$ be the estimation error. The error covariance associated with $\hat{X}_i$ is given by

$$\hat{\Sigma}_i = \text{E}\left\{ X_{e;i} X'_{e;i} | S_{\mathcal{J}_i} \right\}.$$

Applying the results in [4] (page 30) yields the local MMSE estimator and its associated error covariance, specified by

$$\hat{X}_i = \overline{X}_i + \Sigma_{x_i,s_{\mathcal{J}_i}} \Sigma^{+}_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} \left(S_{\mathcal{J}_i} - \overline{S}_{\mathcal{J}_i} \right), \tag{3.1}$$

$$\hat{\Sigma}_i = \Sigma_{x_i,x_i} - \Sigma_{x_i,s_{\mathcal{J}_i}} \Sigma^{+}_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} \Sigma_{s_{\mathcal{J}_i},x_i}, \tag{3.2}$$

respectively, where the superscript of "+" denotes the pseudo-inverse. Although (3.1) provides the local MMSE estimator at node i, its error covariance in (3.2) depends on values of $\{\gamma_{i,j}\}_{j\in\mathcal{N}_i}$ that are random, and unknown in advance. We adopt the Bernoulli model and assume the probability $\mathrm{P}\{\gamma_{i,j}=1\}=p_{i,j}>0$, $j\in\mathcal{N}_i$, with its value determined by the bandwidth and traffic. It follows that at node i and for each $j\in\mathcal{N}_i$,

$$\mathrm{E}\{\gamma_{i,j}\}=p_{i,j},\ \ \mu_{i,j}^2=\mathrm{E}\{(\gamma_{i,j}-p_{i,j})^2\}=p_{i,j}-p_{i,j}^2.$$

A natural question is the minimum average of error covariance $\hat{\Sigma}_i$, denoted by $\widehat{\Sigma}_i$, over $\{\gamma_{i,j}\}_{j\in\mathcal{N}_i}$, and the local estimator, denoted by $\widehat{X}_i$, that minimizes the average error covariance at node i. The answer to this question helps to derive the stationary DKF in the next subsection. Denote $\mathrm{E}_\gamma\{\cdot\}$ as the expectation with respect to $\{\gamma_{i,j}\}$. Then

$$\begin{aligned}\mathrm{E}_\gamma\left\{\begin{bmatrix}\Sigma_{x_i,x_i} & \Sigma_{x_i,s_{\mathcal{J}_i}}\\ \Sigma_{s_{\mathcal{J}_i},x_i} & \Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\end{bmatrix}\right\}&=\mathrm{E}_\gamma\left\{\begin{bmatrix}\Sigma_{x_i,x_i} & \Sigma_{x_i,y_{\mathcal{J}_i}}D_{\gamma;\mathcal{J}_i}\\ D_{\gamma;\mathcal{J}_i}\Sigma_{y_{\mathcal{J}_i},x_i} & D_{\gamma;\mathcal{J}_i}\Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}}D_{\gamma;\mathcal{J}_i}\end{bmatrix}\right\}\\&=\begin{bmatrix}\Sigma_{x_i,x_i} & \Sigma_{x_i,y_{\mathcal{J}_i}}D_{p;\mathcal{J}_i}\\ D_{p;\mathcal{J}_i}\Sigma_{y_{\mathcal{J}_i},x_i} & D_{p;\mathcal{J}_i}\left(W_{\mathcal{J}_i}\circ\Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}}\right)D_{p;\mathcal{J}_i}\end{bmatrix},\end{aligned}\tag{3.3}$$

where, by taking $m_{\mathcal{J}_i}=\sum_{j\in\mathcal{J}_i}m_j$, $p_{i,i}=1$, and $\mu_{i,i}^2=0$,

$$\begin{gathered}D_{p;\mathcal{J}_i}=\mathrm{diag}\{p_{i,j}I_{m_j}\}_{j\in\mathcal{J}_i},\ \ D_{\mu^2;\mathcal{J}_i}=\mathrm{diag}\{\mu_{i,j}^2\mathbf{1}_{m_j}\mathbf{1}_{m_j}'\}_{j\in\mathcal{J}_i},\\ W_{\mathcal{J}_i}=\mathbf{1}_{m_{\mathcal{J}_i}}\mathbf{1}_{m_{\mathcal{J}_i}}'+D_{p;\mathcal{J}_i}^{-1}D_{\mu^2;\mathcal{J}_i}D_{p;\mathcal{J}_i}^{-1}.\end{gathered}$$

Since X and $S_{\mathcal{J}_i}=D_{\gamma;\mathcal{J}_i}Y_{\mathcal{J}_i}$ are not jointly Gaussian in general, we seek the local LMMSE estimator among all linear estimators to minimize the average estimation error variance at sensor node i. The following provides the answer.

Lemma 7. *Among all the unbiased linear estimators, the local estimator that minimizes the average error variance at sensor node i, and its associated minimum average error covariance*

are given respectively by

$$\widehat{X}_i = \overline{X}_i + \Sigma_{x_i,y_{\mathcal{J}_i}} \left(W_{\mathcal{J}_i} \circ \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}} \right)^{-1} D_{p;\mathcal{J}_i}^{-1} \left(S_{\mathcal{J}_i} - D_{\gamma;\mathcal{J}_i} \overline{Y}_{\mathcal{J}_i} \right), \tag{3.4}$$

$$\widehat{\Sigma}_i = \Sigma_{x_i,x_i} - \Sigma_{x_i,y_{\mathcal{J}_i}} \left(W_{\mathcal{J}_i} \circ \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}} \right)^{-1} \Sigma_{y_{\mathcal{J}_i},x_i}. \tag{3.5}$$

Proof. The local LMMSE estimator at node i has the form

$$\widehat{X}_i = G_i S_{\mathcal{J}_i} + h_i, \quad G_i \in \mathbb{R}^{n \times m_{\mathcal{J}_i}}, \quad h_i \in \mathbb{R}^n. \tag{3.6}$$

Setting expectation $\mathrm{E}\{X - \widehat{X}_i\} = 0$ yields

$$h_i = \overline{X}_i - G_i \overline{S}_{\mathcal{J}_i}. \tag{3.7}$$

Thus $\widehat{X}_i = \overline{X}_i + G_i \left(S_{\mathcal{J}_i} - \overline{S}_{\mathcal{J}_i} \right)$, and error covariance

$$\hat{\Sigma}_i = \mathrm{E}\{(X_i - \widehat{X}_i)(X_i - \widehat{X}_i)'\} = \Sigma_{x_i,x_i} + G_i \Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} G_i' - G_i \Sigma_{s_{\mathcal{J}_i},x_i} - \Sigma_{x_i,s_{\mathcal{J}_i}} G_i'.$$

Taking expectation to the above with respect to $\{\gamma_{i,j}\}_{j \in \mathcal{N}_i}$ yields the average error covariance

$$\begin{aligned}
\widehat{\Sigma}_i &= \mathrm{E}_\gamma\{\hat{\Sigma}_i\} = \mathrm{E}_\gamma\{\Sigma_{x_i,x_i}\} + G_i \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\} G_i' - G_i \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},x_i}\} - \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\} G_i' \\
&= \left(G_i - \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\} \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1} \right) \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\} \left(G_i - \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\} \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1} \right)' \\
&\quad + \mathrm{E}_\gamma\{\Sigma_{x_i,x_i}\} - \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\} \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1} \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},x_i}\}.
\end{aligned} \tag{3.8}$$

The variance, i.e., $\mathrm{Tr}\{\widehat{\Sigma}_i\}$ with $\mathrm{Tr}\{\cdot\}$ denoting the trace operation, is minimized by taking

$$G_i = \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\} \mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1}.$$

Substituting the expression above and that in (3.7) into (3.6) and (3.8) leads to

$$\widehat{X}_i = \overline{X}_i + \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\}\mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1}\left(S_{\mathcal{J}_i} - \overline{S}_{\mathcal{J}_i}\right), \tag{3.9}$$

$$\widehat{\Sigma}_i = \mathrm{E}_\gamma\{\Sigma_{x_i,x_i}\} - \mathrm{E}_\gamma\{\Sigma_{x_i,s_{\mathcal{J}_i}}\}\mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}}\}^{-1}\mathrm{E}_\gamma\{\Sigma_{s_{\mathcal{J}_i},x_i}\}, \tag{3.10}$$

which verify the local LMMSE estimator at node i in (3.4) and the associated minimum average error covariance in (3.5), in light of the expressions in (3.3). □

Remark 3. Let $Z = FX$ with F a fixed matrix. Then the local LMMSE estimate of Z at node i and its associated average error covariance can be obtained by modifying (3.4) and (3.5), respectively, yielding

$$\widehat{Z}_i = F\overline{X}_i + F\Sigma_{x_i,y_{\mathcal{J}_i}}\left(W_{\mathcal{J}_i} \circ \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}}\right)^{-1} D^{-1}_{p;\mathcal{J}_i}\left(S_{\mathcal{J}_i} - D_{\gamma;\mathcal{J}_i}\overline{Y}_{\mathcal{J}_i}\right),$$
$$\widehat{\Sigma}_{z;i} = F\Sigma_{x_i,x_i}F' - F\Sigma_{x_i,y_{\mathcal{J}_i}}\left(W_{\mathcal{J}_i} \circ \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}}\right)^{-1}\Sigma_{y_{\mathcal{J}_i},x_i}F'.$$

In the case when $W_{\mathcal{J}_i} \circ \Sigma_{y_{\mathcal{J}_i},y_{\mathcal{J}_i}}$ is singular, pseudo-inverse can be used in (3.4) and (3.5). □

3.1.2 Stationary Distributed Kalman Filter

Data fusion Kalman filter for networked system (2.26) is proposed in Subsection 2.2.2. In this subsection we derive the stationary DKF in the presence of data packet drops for the same networked system described in (2.26). The system dynamics, noise statistics, and initial conditions are all identical to those specified in the beginning of Subsection 2.2.2, thus are not repeated here. In addition, Assumption 5 still holds. Our objective is to design a local Kalman filter at each sensor node so that, via the exchange of sensor measurements among neighboring nodes, the WSN provides a set of optimal local state estimates of the target system. Such a network of local Kalman filters is usually referred to as distributed Kalman filter.

Due to the nature of WSNs, the signal received at the ith sensor node from its neighbor j is given by

$$s_{i,j}(k) = \gamma_{i,j}(k) y_j(k), \quad i \in \mathcal{N}, \quad j \in \mathcal{N}_i, \tag{3.11}$$

where $\gamma_{i,j}(k)$ is imposed to represent the presence of data packet drops. Again the Bernoulli model is adopted, assuming i.i.d. stationary process for $\{\gamma_{i,j}(k)\}$ with probability

$$\mathrm{P}\{\gamma_{i,j}(k) = 1\} = p_{i,j} > 0 \ \ \forall \ k \geq 0.$$

It follows that

$$\mathrm{E}\{\gamma_{i,j}(k)\} = p_{i,j}, \quad \mu_{i,j}^2 = \mathrm{E}\{[\gamma_{i,j}(k) - p_{i,j}]^2\} = p_{i,j} - p_{i,j}^2.$$

We assume that $\gamma_{i,j}(k)$ and $\gamma_{j,i}(k)$ are independent of each other $\forall\ i \neq j$, but $p_{i,j} = p_{j,i}$ is allowed. In addition, $\gamma_{i,j}(k)$ is independent of the noise and the initial state. Based on (3.11), measurement $y_j(k)$ reaches node i at time k if $\gamma_{i,j}(k) = 1$; otherwise all the components of $y_j(k)$ are lost. Denote

$$y_{\mathcal{J}_i}(k) = \mathrm{vec}\{y_j(k)\}_{j\in\mathcal{J}_i}, \quad v_{\mathcal{J}_i}(k) = \mathrm{vec}\{v_j(k)\}_{j\in\mathcal{J}_i}, \quad s_{\mathcal{J}_i}(k) = \mathrm{vec}\{s_{i,j}(k)\}_{j\in\mathcal{J}_i},$$
$$C_{\mathcal{J}_i} = \mathrm{col}\{C_j\}_{j\in\mathcal{J}_i}, \quad R_{\mathcal{J}_i} = \mathrm{diag}\{R_j\}_{j\in\mathcal{J}_i}, \quad D_{\gamma;\mathcal{J}_i}(k) = \mathrm{diag}\{\gamma_{i,j}(k) I_{m_j}\}_{j\in\mathcal{J}_i}, \quad \gamma_{i,i}(k) = 1.$$

Then the state space model at the ith sensor node described in (2.26) and (3.11) can be written in a compact form as

$$x(k+1) = Ax(k) + w(k), \tag{3.12a}$$
$$y_{\mathcal{J}_i}(k) = C_{\mathcal{J}_i} x(k) + v_{\mathcal{J}_i}(k), \tag{3.12b}$$
$$s_{\mathcal{J}_i}(k) = D_{\gamma;\mathcal{J}_i}(k) y_{\mathcal{J}_i}(k). \tag{3.12c}$$

Denote $\mathcal{S}_{\mathcal{J}_i}(k) = \{s_{\mathcal{J}_i}(t)\}_{t=0}^{k}$. Let $\hat{x}_i(k)$ be the MMSE estimate of $x(k)$, and $\Sigma_i(k)$ be the corresponding error covariance at node i, conditioned on observations $\mathcal{S}_{\mathcal{J}_i}(k-1)$, respectively, i.e., $\hat{x}_i(k) = \mathrm{E}[x(k)|\mathcal{S}_{\mathcal{J}_i}(k-1)]$, $\Sigma_i(k) = \mathrm{Cov}[x(k)|\mathcal{S}_{\mathcal{J}_i}(k-1)]$. By convention, $\hat{x}_i(0) = \overline{x}_0$ and $\Sigma_i(0) = \Sigma_0$. Similarly let $\hat{x}_i(k|k)$ be the MMSE estimate of $x(k)$, and $\Sigma_i(k|k)$ be the corresponding error covariance at node i, conditioned on observations $\mathcal{S}_{\mathcal{J}_i}(k)$, respectively, i.e., $\hat{x}_i(k|k) = \mathrm{E}[x(k)|\mathcal{S}_{\mathcal{J}_i}(k)]$, $\Sigma_i(k|k) = \mathrm{Cov}[x(k)|\mathcal{S}_{\mathcal{J}_i}(k)]$. Set $X = x(k)$ and $S_{\mathcal{J}_i} = s_{\mathcal{J}_i}(k) = D_{\gamma;\mathcal{J}_i}(k)y_{\mathcal{J}_i}(k)$. Then direct calculations show that

$$\begin{aligned}
\begin{bmatrix} \overline{X}_i \\ \overline{S}_{\mathcal{J}_i} \end{bmatrix} &= \mathrm{E}\left\{ \begin{bmatrix} x(k) \\ s_{\mathcal{J}_i}(k) \end{bmatrix} \middle| \mathcal{S}_{\mathcal{J}_i}(k-1) \right\} \\
&= \mathrm{E}\left\{ \begin{bmatrix} x(k) \\ D_{\gamma;\mathcal{J}_i}(k)[C_{\mathcal{J}_i}x(k) + v_{\mathcal{J}_i}(k)] \end{bmatrix} \middle| \mathcal{S}_{\mathcal{J}_i}(k-1) \right\} = \begin{bmatrix} \hat{x}_i(k) \\ D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\hat{x}_i(k) \end{bmatrix}, \\
\begin{bmatrix} \Sigma_{x_i,x_i} & \Sigma_{x_i,s_{\mathcal{J}_i}} \\ \Sigma_{s_{\mathcal{J}_i},x_i} & \Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} \end{bmatrix} &= \mathrm{Cov}\left\{ \begin{bmatrix} x(k) \\ s_{\mathcal{J}_i}(k) \end{bmatrix} \middle| \mathcal{S}_{\mathcal{J}_i}(k-1) \right\} \\
&= \begin{bmatrix} \Sigma_i(k) & \Sigma_i(k)C'_{\mathcal{J}_i}D_{\gamma;\mathcal{J}_i}(k) \\ D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\Sigma_i(k) & D_{\gamma;\mathcal{J}_i}(k)[R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\Sigma_i(k)C'_{\mathcal{J}_i}]D_{\gamma;\mathcal{J}_i}(k) \end{bmatrix}, \quad (3.13) \\
\hat{X}_i &= \mathrm{E}\left\{x(k)|\mathcal{S}_{\mathcal{J}_i}(k-1), s_{\mathcal{J}_i}(k)\right\} = \mathrm{E}\left\{x(k)|\mathcal{S}_{\mathcal{J}_i}(k)\right\} = \hat{x}_i(k|k), \\
\hat{\Sigma}_i &= \mathrm{Cov}\left\{x(k)|\mathcal{S}_{\mathcal{J}_i}(k-1), s_{\mathcal{J}_i}(k)\right\} = \mathrm{Cov}\left\{x(k)|\mathcal{S}_{\mathcal{J}_i}(k)\right\} = \Sigma_i(k|k),
\end{aligned}$$

Substituting the above expressions into (3.1) and (3.2) yields the measurement update part of the time-varying local Kalman filter at node i:

$$\begin{aligned}
\hat{x}_i(k|k) &= \hat{x}_i(k) + G_i(k)\left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\hat{x}_i(k)\right], \\
G_i(k) &= \Sigma_i(k)C'_{\mathcal{J}_i}D_{\gamma;\mathcal{J}_i}(k)\left\{D_{\gamma;\mathcal{J}_i}(k)[R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\Sigma_i(k)C'_{\mathcal{J}_i}]D_{\gamma;\mathcal{J}_i}(k)\right\}^{+}, \\
\Sigma_i(k|k) &= \Sigma_i(k) - G_i(k)D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\Sigma_i(k).
\end{aligned}$$

In addition, the time update part of the time-varying local Kalman filter is given by

$$\hat{x}_i(k+1) = A\hat{x}_i(k|k),$$
$$\Sigma_i(k+1) = A\Sigma_i(k|k)A' + Q,$$

by the independence of $w(k)$ and $v_j(k)\ \forall\ j \in \mathcal{J}_i$. Notice that $\{\gamma_{i,j}(k)\}_{j\in\mathcal{N}_i}$ are known and used at time $k+1$ under the TCP-like protocol. Combining the measurement update and the time update yields the time-varying local Kalman filter that is the optimal one-step predictor at node i:

$$\begin{aligned}
\hat{x}_i(k+1) &= A\hat{x}_i(k) + K_i(k)\left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\hat{x}_i(k)\right], \\
K_i(k) &= A\Sigma_i(k)C'_{\mathcal{J}_i}D_{\gamma;\mathcal{J}_i}(k)\left\{D_{\gamma;\mathcal{J}_i}(k)[R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\Sigma_i(k)C'_{\mathcal{J}_i}]D_{\gamma;\mathcal{J}_i}(k)\right\}^{+}, \\
\Sigma_i(k+1) &= A\Sigma_i(k)A' + Q - K_i(k)D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\Sigma_i(k)A'.
\end{aligned}$$

It is important to observe that the above time-varying local Kalman filter applies to the case of time-varying A, Q, and $\{C_i, R_i\}_{i=1}^N$ by adding time index k to the right hand sides. However the random nature of the data packet drops is not taken into consideration. We are interested next in deriving the stationary local Kalman filter that minimizes the local average error variance over the Bernoulli processes $\{\gamma_{i,j}(k)\}$ in the steady-state at each local node. Denote $\mathcal{D}_{\gamma;\mathcal{J}_i}(k) = \{D_{\gamma;\mathcal{J}_i}(t)\}_{t=0}^k$. Let $\widehat{x}_i(k)$ be the LMMSE estimate of $x(k)$ at node i, conditioned on $\mathcal{S}_{\mathcal{J}_i}(k-1)$ and averaged over $\mathcal{D}_{\gamma;\mathcal{J}_i}(k-1)$, and $\widehat{\Sigma}_i(k)$ be the average error covariance associated with $\widehat{x}_i(k)$. Similarly let $\widehat{x}_i(k|k)$ be the LMMSE estimate of $x(k)$ at node i, conditioned on $\mathcal{S}_{\mathcal{J}_i}(k)$ and averaged over $\mathcal{D}_{\gamma;\mathcal{J}_i}(k)$, and $\widehat{\Sigma}_i(k|k)$ be the average error covariance associated with $\widehat{x}_i(k|k)$. The stationary DKF is presented in the following theorem.

Theorem 4. *For the system dynamics described in* (3.12)*, assume that there exists a stabilizing solution* Σ_i *to MARE*

$$\Sigma_i = A\Sigma_i A' + Q - A\Sigma_i C'_{\mathcal{J}_i} \left[W_{\mathcal{J}_i} \circ (R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \Sigma_i C'_{\mathcal{J}_i}) \right]^{-1} C_{\mathcal{J}_i} \Sigma_i A', \tag{3.14}$$

where, by letting $D_{p;\mathcal{J}_i}$ *and* $D_{\mu^2;\mathcal{J}_i}$ *remain the same forms as specified in* (3.3)*,*

$$W_{\mathcal{J}_i} = \mathbf{1}_{m_{\mathcal{J}_i}} \mathbf{1}'_{m_{\mathcal{J}_i}} + D^{-1}_{p;\mathcal{J}_i} D_{\mu^2;\mathcal{J}_i} D^{-1}_{p;\mathcal{J}_i}.$$

The stationary local Kalman filter that minimizes the local average error variance at sensor node i *is specified by*

$$\widehat{x}_i(k+1) = A\widehat{x}_i(k) + K_i \left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k) C_{\mathcal{J}_i} \widehat{x}_i(k) \right], \tag{3.15a}$$

$$K_i = A\Sigma_i C'_{\mathcal{J}_i} \left[W_{\mathcal{J}_i} \circ (R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \Sigma_i C'_{\mathcal{J}_i}) \right]^{-1} D^{-1}_{p;\mathcal{J}_i}. \tag{3.15b}$$

Proof. Set $X = x(k)$ and $S_{\mathcal{J}_i} = s_{\mathcal{J}_i}(k) = D_{\gamma;\mathcal{J}_i}(k) y_{\mathcal{J}_i}(k)$. Recall the expressions in (3.13). Straightforward calculations yield

$$\begin{aligned}
&\mathrm{E}_\gamma \left\{ \begin{bmatrix} \Sigma_{x_i,x_i} & \Sigma_{x_i,s_{\mathcal{J}_i}} \\ \Sigma_{s_{\mathcal{J}_i},x_i} & \Sigma_{s_{\mathcal{J}_i},s_{\mathcal{J}_i}} \end{bmatrix} \right\} \\
&\quad = \mathrm{E}_\gamma \left\{ \begin{bmatrix} \Sigma_i(k) & \Sigma_i(k) C'_{\mathcal{J}_i} D_{\gamma;\mathcal{J}_i}(k) \\ D_{\gamma;\mathcal{J}_i}(k) C_{\mathcal{J}_i} \Sigma_i(k) & D_{\gamma;\mathcal{J}_i}(k) [R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \Sigma_i(k) C'_{\mathcal{J}_i}] D_{\gamma;\mathcal{J}_i}(k) \end{bmatrix} \right\} \\
&\quad = \begin{bmatrix} \widehat{\Sigma}_i(k) & \widehat{\Sigma}_i(k) C'_{\mathcal{J}_i} D_{p;\mathcal{J}_i} \\ D_{p;\mathcal{J}_i} C_{\mathcal{J}_i} \widehat{\Sigma}_i(k) & D_{p;\mathcal{J}_i} \left\{ W_{\mathcal{J}_i} \circ [R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \widehat{\Sigma}_i(k) C'_{\mathcal{J}_i}] \right\} D_{p;\mathcal{J}_i} \end{bmatrix}.
\end{aligned} \tag{3.16}$$

In addition, there hold

$$\overline{X}_i = \widehat{x}_i(k), \;\; \overline{S}_{\mathcal{J}_i} = D_{\gamma;\mathcal{J}_i}(k) C_{\mathcal{J}_i} \widehat{x}_i(k), \;\; \widehat{X}_i = \widehat{x}_i(k|k), \;\; \widehat{\Sigma}_i = \widehat{\Sigma}_i(k|k).$$

Substituting the expressions above and those in (3.16) into (3.9) and (3.10) yields the measurement update part of the time-varying local Kalman filter at node i:

$$\begin{aligned}
\widehat{x}_i(k|k) &= \widehat{x}_i(k) + \widehat{G}_i(k)\left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\widehat{x}_i(k)\right], \\
\widehat{G}_i(k) &= \widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}\left\{W_{\mathcal{J}_i} \circ [R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}]\right\}^{-1} D^{-1}_{p;\mathcal{J}_i}, \\
\widehat{\Sigma}_i(k|k) &= \widehat{\Sigma}_i(k) - \widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}\left\{W_{\mathcal{J}_i} \circ [R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}]\right\}^{-1} C_{\mathcal{J}_i}\widehat{\Sigma}_i(k).
\end{aligned}$$

In light of the independence of $w(k)$ and $v_j(k)\ \forall\ j \in \mathcal{J}_i$, the time update part of the time-varying local Kalman filter is given by

$$\begin{aligned}
\widehat{x}_i(k+1) &= A\widehat{x}_i(k|k), \\
\widehat{\Sigma}_i(k+1) &= A\widehat{\Sigma}_i(k|k)A' + Q.
\end{aligned}$$

Combining the measurement update and the time update yields the time-varying local Kalman filter at node i:

$$\begin{aligned}
\widehat{x}_i(k+1) &= A\widehat{x}_i(k) + \widehat{K}_i(k)\left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\widehat{x}_i(k)\right], \\
\widehat{K}_i(k) &= A\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}\left\{W_{\mathcal{J}_i} \circ [R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}]\right\}^{-1} D^{-1}_{p;\mathcal{J}_i},
\end{aligned}$$

where the minimum local average error covariance follows MDRE

$$\widehat{\Sigma}_i(k+1) = A\widehat{\Sigma}_i(k)A' + Q - A\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}\left\{W_{\mathcal{J}_i} \circ [R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\widehat{\Sigma}_i(k)C'_{\mathcal{J}_i}]\right\}^{-1} C_{\mathcal{J}_i}\widehat{\Sigma}_i(k)A'. \quad (3.17)$$

As $k \to \infty$, MDRE (3.17) converges to MARE (3.14), $\widehat{\Sigma}_i(k) \to \Sigma_i$, and $\widehat{K}_i(k) \to K_i$. The hypothesis implies that the stabilizing solution to MARE (3.14) exists. It follows that (3.15) is indeed the stationary local Kalman filter that minimizes the local average error variance at the ith sensor node. □

Remark 4. If $w(k)$ and $v_j(k)$ are correlated for some $j \in \mathcal{J}_i$, then

$$M_{\mathcal{J}_i} = \mathrm{E}\{v_{\mathcal{J}_i}(k)w(k)'\} \neq 0,$$

in general. In this case (3.12a) can be written as

$$x(k+1) = \tilde{A}_i x(k) + M'_{\mathcal{J}_i} R^{-1}_{\mathcal{J}_i} y_{\mathcal{J}_i}(k) + \tilde{w}_i(k),$$

where $\tilde{A}_i = A - M'_{\mathcal{J}_i} R^{-1}_{\mathcal{J}_i} C_{\mathcal{J}_i}$ and $\tilde{w}_i(k) = w(k) - M'_{\mathcal{J}_i} R^{-1}_{\mathcal{J}_i} v_{\mathcal{J}_i}(k)$. As a result, $\tilde{w}_i(k)$ and $v_{\mathcal{J}_i}(k)$ are uncorrelated, and are thus independent of each other. Recall that $w(k)$ and $\{v_i(k)\}_{i=1}^N$ are all Gaussian distributed. The results in Theorem 4 can be applied to derive the stationary local Kalman filter that has the same form as (3.15a); However the estimation gain K_i and MARE are replaced by

$$\begin{aligned}
K_i &= (A\Sigma_i C'_{\mathcal{J}_i} + M'_{\mathcal{J}_i}) \left[W_{\mathcal{J}_i} \circ (R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \Sigma_i C'_{\mathcal{J}_i}) \right]^{-1} D^{-1}_{p;\mathcal{J}_i}, \\
\Sigma_i &= A\Sigma_i A' + Q - (A\Sigma_i C'_{\mathcal{J}_i} + M'_{\mathcal{J}_i}) \left[W_{\mathcal{J}_i} \circ (R_{\mathcal{J}_i} + C_{\mathcal{J}_i} \Sigma_i C'_{\mathcal{J}_i}) \right]^{-1} (C_{\mathcal{J}_i} \Sigma_i A' + M_{\mathcal{J}_i}),
\end{aligned}$$

respectively. Detailed derivations are skipped. It is noted that if $M_{\mathcal{J}_i} = 0$, the above expressions collapse to those of K_i and MARE in Theorem 4, respectively. □

3.2 Existence of Stabilizing Solution to MARE

In this section we investigate the necessary and sufficient condition for the existence of the stabilizing solution to MARE (3.14), which is hinged on the MS stabilizability of the corresponding estimation error dynamics. Denote the state estimation error at node i by $e_{x;i}(k) = x(k) - \widehat{x}_i(k)$. The estimation error dynamics associated to filter (3.15) are given by

$$e_{x;i}(k+1) = \left[A - K_i D_{\gamma;\mathcal{J}_i}(k) C_{\mathcal{J}_i} \right] e_{x;i}(k), \tag{3.18}$$

with noise terms removed. Specially we adopt

$$D_{\gamma;\mathcal{J}_i}(k) = D_{p;\mathcal{J}_i}\left[I + D_{\delta;\mathcal{J}_i}(k)\right], \quad D_{\delta;\mathcal{J}_i}(k) = \mathrm{diag}\{\delta_{i,j}(k)I_{m_j}\}_{j\in\mathcal{J}_i}, \tag{3.19}$$

where $\{\delta_{i,j}(k)\}_{j\in\mathcal{N}_i}$ are mutually independent white random processes with mean zero and variance

$$\mathrm{E}\{\delta_{i,j}(k)^2\} = \nu_{i,j}^2 = \frac{\mu_{i,j}^2}{p_{i,j}^2} = p_{i,j}^{-1} - 1 \iff p_{i,j} = \frac{1}{1+\nu_{i,j}^2},$$

and $\delta_{i,i}(k) = 0$, $\nu_{i,i}^2 = 0$. Substituting (3.19) into (3.18) yields the error dynamics in feedback form as

$$e_{x;i}(k+1) = (A + K_{p;i}C_{\mathcal{J}_i})e_{x;i}(k) + K_{p;i}D_{\delta;\mathcal{J}_i}(k)e_{y;i}(k), \quad e_{y;i}(k) = C_{\mathcal{J}_i}e_{x;i}(k),$$

where $K_{p;i} = -K_i D_{p;\mathcal{J}_i}$. The above feedback system can be schematically illustrated by the block diagram in Figure 3.1.

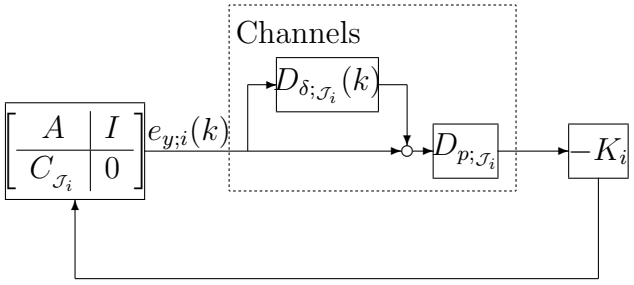


Figure 3.1: Estimation error dynamics of DKF in feedback form

In light of [74, 103], MARE (3.14) admits a unique stabilizing solution if and only if MS stabilizability holds for error dynamics (3.18). Again the stability margin, termed as $\alpha_{i;\max} > 0$, is adopted to help to determine the MS stabilizability. The next result presents an efficient algorithm to compute $\alpha_{i;\max}$.

Theorem 5. *Under Assumption* 5 *and the detectability of* $(C_{\mathcal{J}_i}, A)$*, for a given set of variances* $\{\nu_{i,j}^2\}_{j\in\mathcal{N}_i}$*, the stability margin* $\alpha_{i;\max} = \beta_{i;\min}^{-1}$ *can be computed via the minimization of* β_i

subject to the following LMIs *over matrix pair* $(X_{\Phi;i}^{-1}, \Phi_{\mathcal{J}_i}^{-2})$*:*

$$\begin{aligned}
&\text{(a) } \beta_i \nu_{i,j}^{-2} X_{\Phi;i}^{-1} > C_j' \Phi_j^{-2} C_j, \ \ \forall\, j \in \mathcal{N}_i, \\
&\text{(b) } X_{\Phi;i}^{-1} + C_{\mathcal{J}_i}' \Phi_{\mathcal{J}_i}^{-2} C_{\mathcal{J}_i} \geq A' X_{\Phi;i}^{-1} A, \\
&\text{(c) } X_{\Phi;i}^{-1} > 0, \quad \Phi_{\mathcal{J}_i}^{-2} > 0,
\end{aligned}$$

where $\Phi_{\mathcal{J}_i} = \text{diag}\{\Phi_j\}_{j\in\mathcal{J}_i}$*. Furthermore, error dynamics* (3.18) *are MS stabilizable, if and only if* $\alpha_{i;\max} = \beta_{i;\min}^{-1} \geq 1$*.*

The proof for Theorem 5 is similar to that for Theorem 3 in Subsection 2.2.3, thus is omitted here. The comments on the LMI problem and the stability margin in Remark 2 hold for Theorem 5 as well. Given a set of variances $\{\nu_{i,j}^2\}_{j\in\mathcal{N}_i}$, if $\alpha_{i;\max} \geq 1$, then Under both detectability of $(C_{\mathcal{J}_i}, A)$ and stabilizability of $(A, Q^{1/2})$, the stabilizing solution Σ_i to MARE (3.14) can be obtained by computing $\widehat{\Sigma}_i(k)$ iteratively following MDRE (3.17), i.e., $\Sigma_i = \lim_{k\to\infty} \widehat{\Sigma}_i(k)$. More details can be referred to Subsection 2.1.3.

3.3 Kalman Consensus Filter

The previous two subsections are focused on the stationary DKF, and its MS stability. The LMMSE estimates of these local Kalman filters differ from each other in general. The KCF is proposed in [53, 54, 55] to coordinate the local Kalman filters and achieve the consensus on state estimation among all nodes without taking data packet drops into consideration. It has the advantage that any local node can provide its own state estimate with high reliability; the price to be paid is that each sensor node $j \in \mathcal{N}_i$ needs to send not only its measurement $y_j(k)$, but also its state estimate $\widehat{x}_j(k)$ to node i. In this section, we devise the KCF in the presence of data packet drops as a generalization of the stationary DKF proposed in previous subsections, followed by a formal MS stability analysis of the associated estimation error dynamics.

3.3.1 Preliminaries

Before presenting the main results, we need to introduce several key lemmas that will help address the stability issue. Consider the stochastic feedback system configured in Figure 3.2. The discrete-time stochastic system $S_{\delta(k)}$ is described by

$$\chi(k+1) = \left[A_{s;0} + \sum_{i=1}^{N} A_{s;i}\delta_i(k)\right]\chi(k) + \left[B_{s;0} + \sum_{i=1}^{N} B_{s;i}\delta_i(k)\right]u(k), \tag{3.20a}$$

$$y(k) = C_s\chi(k) + D_s u(k), \tag{3.20b}$$

where $\{\delta_i(k)\}_{i=1}^{N}$ are mutually independent white random processes with mean zero and variances $\{\sigma_i^2\}_{i=1}^{N}$. The $\mathcal{H}_\infty$ norm for stochastic systems is defined next.

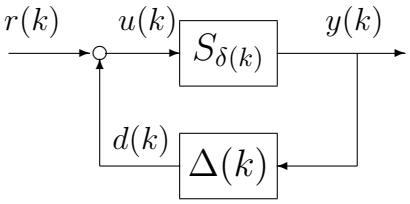


Figure 3.2: Stochastic feedback system

Definition 3. [20, 21] *Suppose that stochastic system $S_{\delta(k)}$ in* (3.20) *is MS stable. There exists a constant $h_\infty \geq 0$ such that*

$$\sqrt{\sum_{k=0}^{T} \mathrm{E}\left\{\|y(k)\|^2\right\}} \leq h_\infty \sqrt{\sum_{k=0}^{T} \mathrm{E}\left\{\|u(k)\|^2\right\}} \quad \forall\, T \geq 0. \tag{3.21}$$

The $\mathcal{H}_\infty$ norm of $S_{\delta(k)}$, denoted by $\|S_{\delta(k)}\|_{\mathcal{H}_\infty}$, is defined as the minimal $h_\infty \geq 0$ such that (3.21) *is satisfied.*

If $S_{\delta(k)}$ is MS stable, then its $\mathcal{H}_\infty$ norm is bounded. Let $h_\infty > 0$ be a strict upper bound for $\|S_{\delta(k)}\|_{\mathcal{H}_\infty}$. Theorem 2.5 of [20] provides the following characterization for the $\mathcal{H}_\infty$ norm, referred to as *Bounded Real Lemma* for discrete-time stochastic systems.

Lemma 8. [20] *System* (3.20) *is MS stable, and* $\|S_{\delta(k)}\|_{\mathcal{H}_\infty} < h_\infty$, *if and only if there exists* $P_s > 0$ *such that*

$$P_s > A'_{s;0}P_sA_{s;0} + C'_sC_s + \Pi'_s\Lambda_s^{-1}\Pi_s + \sum_{i=1}^{N}\sigma_i^2 A'_{s;i}P_sA_{s;i}, \tag{3.22}$$

$$\Lambda_s = h_\infty^2 I - B'_{s;0}P_sB_{s;0} - D'_sD_s - \sum_{i=1}^{N}\sigma_i^2 B'_{s;i}P_sB_{s;i} > 0, \tag{3.23}$$

$$\Pi_s = B'_{s;0}P_sA_{s;0} + D'_sC_s + \sum_{i=1}^{N}\sigma_i^2 B'_{s;i}P_sA_{s;i}.$$

It is important to observe that (3.22) and (3.23) are equivalent to the following LMI [20]:

$$\begin{bmatrix} P_s - A'_{s;0}P_sA_{s;0} - C'_sC_s - \sum\limits_{i=1}^{N}\sigma_i^2 A'_{s;i}P_sA_{s;i} & \Pi'_s \\ \Pi_s & \Lambda_s \end{bmatrix} > 0,$$

in light of *Schur Complement Lemma*.

For the feedback system in Figure 3.2, we assume that $r(k)$ is an arbitrary vector-valued white process with mean zero, and is independent of $\chi(0) = \chi_0$, the initial condition of system (3.20), for all $k \geq 0$. Its auto-covariance satisfies

$$\mathrm{E}\{r(k)r(k-\tau)'\} = \delta_{\mathrm{K}}(\tau)\sigma_r(k)^2 I$$

with $\delta_{\mathrm{K}}(\cdot)$ denoting the Kronecker Delta function, and

$$p_T[r(k)] := \sqrt{\frac{1}{T}\sum_{k=0}^{T-1}\mathrm{E}\{\|r(k)\|^2\}}$$

being bounded for all $T > 0$, including $T = \infty$. The next lemma verifies that the above nonnegative function $p_T(\cdot)$ is a semi-norm [42].

Lemma 9. *Let $\{a(k)\}$ and $\{b(k)\}$ be vector-valued random processes of the same dimension. There holds the triangle inequality*

$$p_T[a(k)+b(k)] \le p_T[a(k)] + p_T[b(k)].$$

Proof. Let $x = [x_1 \ \dots \ x_N]' \in \mathbb{R}^N$ and $y = [y_1 \ \dots \ y_N]' \in \mathbb{R}^N$ be random vectors of the same dimension N. In light of the well-known Cauchy-Schwarz Inequality [81],

$$\mathrm{E}\{x'y\} \le |\mathrm{E}\{x'y\}| \le \sum_{i=1}^{N} |\mathrm{E}\{x_i y_i\}| \le \sum_{i=1}^{N} \sqrt{\mathrm{E}\{x_i^2\}\mathrm{E}\{y_i^2\}} \le \sqrt{\sum_{i=1}^{N} \mathrm{E}\{x_i^2\}} \sqrt{\sum_{i=1}^{N} \mathrm{E}\{y_i^2\}}.$$

Then there holds

$$\mathrm{E}\{\|x+y\|^2\} = \mathrm{E}\{\|x\|^2\} + \mathrm{E}\{\|y\|^2\} + 2\mathrm{E}\{x'y\} \le \mathrm{E}\{\|x\|^2\} + \mathrm{E}\{\|y\|^2\} + 2\sqrt{\mathrm{E}\{\|x\|^2\}}\sqrt{\mathrm{E}\{\|y\|^2\}}.$$

Taking square root on both sides of the above inequality yields

$$\sqrt{\mathrm{E}\{\|x+y\|^2\}} \le \sqrt{\mathrm{E}\{\|x\|^2\}} + \sqrt{\mathrm{E}\{\|y\|^2\}}.$$

Denote $\underline{a} = \mathrm{vec}\{a(0), \dots, a(T-1)\}$, $\underline{b} = \mathrm{vec}\{b(0), \dots, b(T-1)\}$. In light of the above triangle inequality, there holds

$$\begin{aligned} p_T[a(k)+b(k)] &= \sqrt{\frac{1}{T}} \sqrt{\mathrm{E}\left\{\sum_{k=0}^{T-1} \|a(k)+b(k)\|^2\right\}} = \sqrt{\frac{1}{T}} \sqrt{\mathrm{E}\left\{\|\underline{a}+\underline{b}\|^2\right\}} \\ &\le \sqrt{\frac{1}{T}} \sqrt{\mathrm{E}\left\{\|\underline{a}\|^2\right\}} + \sqrt{\frac{1}{T}} \sqrt{\mathrm{E}\left\{\|\underline{b}\|^2\right\}} = p_T[a(k)] + p_T[b(k)], \end{aligned}$$

which concludes the proof. □

The next result is concerned with the MS stability of the feedback system in Figure 3.2.

Lemma 10. *Suppose that system $S_{\delta(k)}$ described in* (3.20) *is MS stable,* $\|S_{\delta(k)}\|_{\mathcal{H}_\infty} < h_\infty$*, and* $\Delta(k)'\Delta(k) \leq h_\infty^{-2} I$ *for all* $k \geq 0$*, where* $h_\infty > 0$*. Then the feedback system in Figure* 3.2 *is MS stable.*

Proof. In light of the results in [20], the hypothesis on $\|\mathcal{S}_{\delta(k)}\|_{\mathcal{H}_\infty} < h_\infty$ implies the existence of $\epsilon > 0$ such that for each integer $T > 0$, there holds inequality

$$\sum_{k=0}^{T-1} \mathrm{E}\{\|y(k)\|^2\} \leq (h_\infty^2 - \epsilon^2) \sum_{k=0}^{T-1} \mathrm{E}\{\|u(k)\|^2\},$$

which is equivalent to

$$p_T[y(k)] \leq \sqrt{h_\infty^2 - \epsilon^2}\, p_T[u(k)], \tag{3.24}$$

by the definition of $p_T(\cdot)$. According to Figure 3.2, at each time index k, there hold $u(k) = r(k) + d(k)$ and $d(k) = \Delta(k)y(k)$. Since $p_T(\cdot)$ is a semi-norm by Lemma 9, there holds

$$p_T[u(k)] = p_T[r(k) + \Delta(k)y(k)] \leq p_T[r(k)] + p_T[\Delta(k)y(k)] \leq p_T[r(k)] + h_\infty^{-1} p_T[y(k)].$$

Substituting (3.24) into the above triangle inequality yields

$$p_T[u(k)] \leq p_T[r(k)] + \sqrt{1 - \epsilon^2/h_\infty^2}\, p_T[u(k)] \leq p_T[r(k)] / \left(1 - \sqrt{1 - \epsilon^2/h_\infty^2}\right)$$

that is bounded for each $T > 0$. Therefore $p_T[y(k)]$ is also bounded for each $T > 0$ by (3.24). Taking the limit $T \to \infty$ unveils the boundedness of

$$\lim_{T\to\infty} p_T[u(k)] = \lim_{k\to\infty} \sqrt{\mathrm{E}\{\|u(k)\|^2\}} \quad \text{and} \quad \lim_{T\to\infty} p_T[y(k)] = \lim_{k\to\infty} \sqrt{\mathrm{E}\{\|y(k)\|^2\}},$$

thereby concluding the MS stability of the closed-loop system in Figure 3.2. □

3.3.2 Main Results

Given the same networked system model, noise statistics, and initial conditions as specified in Subsection 3.1.2, we now include the exchange of prior estimates into the design of local Kalman filter at node i, $i \in \mathcal{N}$ as

$$\widehat{x}_i(k+1) = A\widehat{x}_i(k) + K_i\left[s_{\mathcal{J}_i}(k) - D_{\gamma;\mathcal{J}_i}(k)C_{\mathcal{J}_i}\widehat{x}_i(k)\right] + \varepsilon\sum_{j\in\mathcal{N}_i}\gamma_{i,j}(k)A\left[\widehat{x}_j(k) - \widehat{x}_i(k)\right], \tag{3.25a}$$

$$K_i = A\Sigma_i C'_{\mathcal{J}_i}\left[W_{\mathcal{J}_i} \circ (R_{\mathcal{J}_i} + C_{\mathcal{J}_i}\Sigma_i C'_{\mathcal{J}_i})\right]^{-1} D^{-1}_{p;\mathcal{J}_i}, \tag{3.25b}$$

where Σ_i is the stabilizing solution to MARE (3.14), and consensus coefficient $\varepsilon > 0$.

We assume that both sensor measurement $y_j(k)$ and prior estimate $\widehat{x}_j(k)$ are sent from node j to its neighboring node i at time instant k, through the same communication channel. Hence the aforementioned $\gamma_{i,j}(k)$ is also imposed in the newly-added estimate term to indicate the occurrence of packet drops in the transmission of prior estimate information. Although $\{\widehat{x}_j(k)\}_{j\in\mathcal{N}_i}$ may help to improve the local estimation at sensor node i, they are primarily used to achieve the consensus of state estimation from N different sensor nodes, following the pioneering work in [53, 54, 55]. We call the above estimator Kalman consensus filter, in accordance of [55]. Obviously, KCF (3.25) consists of two parts: the stationary DKF, and an additional consensus term on the difference between local prior estimate and those from neighboring nodes. Intuitively, adopting the consensus term can help to reduce the disagreement among different local estimates.

Let

$$\mathcal{L}_\gamma(k) = \left[l_{i,j}(k)\right], \quad l_{i,j}(k) = \begin{cases} \sum\limits_{j\in\mathcal{N}_i}\gamma_{i,j}(k), & j = i, \\ -\gamma_{i,j}(k), & j \in \mathcal{N}_i, \\ 0, & j \notin \mathcal{J}_i, \end{cases}$$

be the stochastic Laplacian matrix associated with the stochastic time-varying graph $\mathcal{G}(k)$

derived from the sensor network in the presence of packet drops. Note that based on the configuration of our sensor network of N nodes, there are only a finite number of possible communication graphs, say a total of φ. Denote the set of all possible graphs by $\{\mathcal{G}_1, \ldots, \mathcal{G}_\varphi\}$, and the set of corresponding Laplacian matrices by $\{\mathcal{L}_1, \ldots, \mathcal{L}_\varphi\}$, i.e., $\mathcal{G}(k) \in \{\mathcal{G}_1, \ldots, \mathcal{G}_\varphi\}$, $\mathcal{L}_\gamma(k) \in \{\mathcal{L}_1, \ldots, \mathcal{L}_\varphi\}$, $\forall\, k \geq 0$. Then we can define

$$\xi^\star := \max\left\{\overline{\sigma}\left(\mathcal{L}_i\right)\right\}_{i=1}^{\varphi} \geq \max\left\{\overline{\sigma}\left[\mathcal{L}_\gamma(k)\right]\right\}_{k\geq 0} > 0 \tag{3.26}$$

with $\overline{\sigma}\{\cdot\}$ denoting the largest singular value.

The estimation error dynamics associated to KCF (3.25) at node i are given by

$$e_{x;i}(k+1) = \left[A - K_i D_{\gamma;\mathcal{J}_i}(k) C_{\mathcal{J}_i}\right] e_{x;i}(k) + \varepsilon \sum_{j\in\mathcal{N}_i} \gamma_{i,j}(k) A\left[e_{x;j}(k) - e_{x;i}(k)\right], \tag{3.27}$$

with noise terms removed. Define $e_x(k) := \mathrm{vec}\{e_{x;1}(k), \ldots, e_{x;N}(k)\}$. For convenience, denote $\widetilde{\mathcal{L}}_\gamma(k) = \mathcal{L}_\gamma(k) \otimes I_n$ and $\widetilde{A}_K = \widetilde{A} - \widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{C}_{\mathcal{J}}$, where

$$\begin{aligned} \widetilde{A} &= \mathrm{diag}(A, \ldots, A), \qquad & \widetilde{C}_{\mathcal{J}} &= \mathrm{diag}(C_{\mathcal{J}_1}, \ldots, C_{\mathcal{J}_N}), \\ \widetilde{K} &= \mathrm{diag}(K_1, \ldots, K_N), & \widetilde{D}_{p;\mathcal{J}} &= \mathrm{diag}(D_{p;\mathcal{J}_1}, \ldots, D_{p;\mathcal{J}_N}). \end{aligned}$$

Furthermore, with $\widetilde{D}_{\delta;\mathcal{J}}(k) = \mathrm{diag}\{D_{\delta;\mathcal{J}_1}(k), \ldots, D_{\delta;\mathcal{J}_N}(k)\}$, the collective model for error dynamics (3.27) can be written in a compact form as

$$e_x(k+1) = \left[\widetilde{A}_K - \widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{D}_{\delta;\mathcal{J}}(k)\widetilde{C}_{\mathcal{J}}\right] e_x(k) - \varepsilon\widetilde{\mathcal{L}}_\gamma(k)\widetilde{A}e_x(k). \tag{3.28}$$

The above is equivalent to a feedback system in which the forward path is described by

$$e_x(k+1) = \left[\widetilde{A}_K - \widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{D}_{\delta;\mathcal{J}}(k)\widetilde{C}_{\mathcal{J}}\right] e_x(k) + u_\varepsilon(k), \;\; e_\varepsilon(k) = \widetilde{A}e_x(k) \tag{3.29}$$

with input $u_\varepsilon(k)$ and output $e_\varepsilon(k)$, and the feedback path is described by

$$u_\varepsilon(k) = -\varepsilon\widetilde{\mathcal{L}}_\gamma(k)e_\varepsilon(k). \tag{3.30}$$

Moreover, we can rewrite the term $\widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{D}_{\delta;\mathcal{J}}(k)\widetilde{C}_{\mathcal{J}}$ in (3.29) in a summation form as

$$\widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{D}_{\delta;\mathcal{J}}(k)\widetilde{C}_{\mathcal{J}} = \sum_{i=1}^{N}\sum_{j\in\mathcal{J}_i} p_{i,j}\widetilde{K}_{(i,j)}\widetilde{C}_{\mathcal{J}}^{(i,j)}\delta_{i,j}(k)$$

with $\widetilde{K}_{(i,j)}$ being the columns of $\widetilde{K}$ associated to $\delta_{i,j}(k)$, and $\widetilde{C}_{\mathcal{J}}^{(i,j)}$ being the rows of $\widetilde{C}_{\mathcal{J}}$ associated to $\delta_{i,j}(k)$. The sufficient condition for the MS stability of error dynamics (3.28) is demonstrated in the following result.

Theorem 6. *Suppose that $(C_{\mathcal{J}_i}, A)$ is detectable, Assumption 5 holds, and the error dynamics described in (3.18) are MS stabilizable $\forall\ i \in \mathcal{N}$. Then MS stability holds for the error dynamics (3.28) associated to KCF (3.25) $\forall\ i \in \mathcal{N}$, if the consensus coefficient ε in (3.28) satisfies $0 < \varepsilon \leq (\xi^\star h_\infty^\star)^{-1}$, where $\xi^\star$ is defined in (3.26), and $h_\infty^\star$ can be computed via the minimization of $h_\infty > 0$ subject to the following LMI over $P > 0$:*

$$\begin{bmatrix} P - \widetilde{A}_K' P \widetilde{A}_K - \widetilde{A}'\widetilde{A} - M & \widetilde{A}_K' P \\ P\widetilde{A}_K & h_\infty^2 I - P \end{bmatrix} > 0 \tag{3.31}$$

with $M = \sum_{i=1}^{N}\sum_{j\in\mathcal{J}_i} \mu_{i,j}^2 \left[\widetilde{K}_{(i,j)}\widetilde{C}_{\mathcal{J}}^{(i,j)}\right]' P \widetilde{K}_{(i,j)}\widetilde{C}_{\mathcal{J}}^{(i,j)}$.

Proof. The hypotheses in Theorem 6 guarantee the existence of the stabilizing solution to MARE (3.14) and the optimal estimation gain in (3.25b). Then by continuity argument, MS stability holds for error dynamics (3.28) when ε is sufficiently small. By *Schur Complement Lemma*, LMI (3.31) is equivalent to $\Lambda = h_\infty^2 I - P > 0$ and

$$P > \widetilde{A}_K' P \widetilde{A}_K + \widetilde{A}'\widetilde{A} + M + \widetilde{A}_K' P \Lambda^{-1} P \widetilde{A}_K. \tag{3.32}$$

In light of Lemma 8, $h_\infty^\star$ being the minimum of h_∞ such that $\Lambda = h_\infty^2 I - P > 0$ and inequality (3.32) holds implies that the $\mathcal{H}_\infty$ norm of system $S_{\delta(k)}$ in (3.20) with

$$A_{s;0} = \widetilde{A}_K, \quad \sum_{i=1}^{N} A_{s;i}\delta_i(k) = -\widetilde{K}\widetilde{D}_{p;\mathcal{J}}\widetilde{D}_{\delta;\mathcal{J}}(k)\widetilde{C}_{\mathcal{J}}, \quad B_{s;0} = I, \quad B_{s;i} = 0, \quad C_s = \widetilde{A}, \quad D_s = 0,$$

is strictly upper bounded by $h_\infty^\star$. By the property of Kronecker product [9], we have

$$\overline{\sigma}\left[\widetilde{\mathcal{L}}_\gamma(k)\right] = \overline{\sigma}\left[\mathcal{L}_\gamma(k)\right].$$

Thus, given $0 < \varepsilon \leq (\xi^\star h_\infty^\star)^{-1}$, there holds

$$\varepsilon^2 \widetilde{\mathcal{L}}_\gamma(k)' \widetilde{\mathcal{L}}_\gamma(k) \leq (\xi^\star h_\infty^\star)^{-2}(\xi^\star)^2 I = (h_\infty^\star)^{-2} I$$

for all $k \geq 0$. Although $\widetilde{\mathcal{L}}_\gamma(k)$ is random, the finiteness of different $\widetilde{\mathcal{L}}_\gamma(k)$ allows computing the upper bound before taking expectation for each of its entries. Hence the above inequality implies the MS stability of feedback system described by (3.29) and (3.30) in light of Lemma 10 with $\Delta(k) = -\varepsilon\widetilde{\mathcal{L}}_\gamma(k)$, thereby concluding the proof. □

While Theorem 6 provides an elegant way to compute the upper bound for the consensus coefficient ε, the dimension and consequently the computation complexity of LMI (3.31) can be quite high, especially when N, the total number of sensor nodes, is very large. Building on Theorem 6, we consider P as a block diagonal matrix to decouple LMI (3.31). The result is presented in the following corollary.

Corollary 1. *Assume that the hypotheses in Theorem* 6 *hold. Then MS stability holds for the error dynamics* (3.28) *associated to KCF* (3.25) $\forall\ i \in \mathcal{N}$*, if the consensus coefficient* ε *in* (3.28) *satisfies* $0 < \varepsilon \leq (\xi^\star h_\infty^\star)^{-1}$*, where* $\xi^\star$ *is defined in* (3.26)*, and* $h_\infty^\star$ *can be computed via*

the minimization of $h_\infty > 0$ *subject to the following LMIs over* $P = \mathrm{diag}(P_1, \ldots, P_N) > 0$*:*

$$\begin{bmatrix} P_i - A'_{K;i} P_i A_{K;i} - A'A - M_i & A'_{K;i} P_i \\ P_i A_{K;i} & h_\infty^2 I - P_i \end{bmatrix} > 0, \ \ \forall\, i \in \mathcal{N} \tag{3.33}$$

with $A_{K;i} = A - K_i D_{p;\mathcal{J}_i} C_{\mathcal{J}_i}$ *and* $M_i = C'_{\mathcal{J}_i} \left[D_{\mu^2;\mathcal{J}_i} \circ (K_i' P_i K_i) \right] C_{\mathcal{J}_i}$.

Proof. By *Schur Complement Lemma*, LMIs (3.33) are equivalent to

$$P_i > A'_{K;i} P_i A_{K;i} + A'A + M_i + A'_{K;i} P_i \Lambda_i^{-1} P_i A_{K;i},$$
$$\Lambda_i = h_\infty^2 I - P_i > 0, \ \ \forall\, i \in \mathcal{N},$$

which can be written in a compact form as (3.32) and

$$\Lambda = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_N) = h_\infty^2 I - P > 0,$$

respectively. The rest of the proof is the same as that for Theorem 6, thus is omitted here. □

Since the minimum h_∞ may be achieved when matrix P is not block diagonal, the upper bound for ε obtained by Corollary 1 is usually more conservative than the one obtained by Theorem 6. However the algorithm in Corollary 1 shows its great advantage in computation complexity for large-scale WSNs due to the decoupled LMIs. In addition, although Theorem 6 and Corollary 1 provide elegant analytical methods to find the upper bound for ε, the results can be quite conservative since the sufficient condition considered in the derivations may not be necessary for the MS stability of the error dynamics. In other words, consensus coefficients greater than the upper bound obtained by Theorem 6 or Corollary 1 may still work well for KCF (3.25).

Theorem 6 and Corollary 1 provide the guidelines for implementing KCF (3.25). First the stability margin $\alpha_{i;\max}$ can be computed by solving the LMI problem in Theorem 5 to

determine the MS stabilizability of the error dynamics described in (3.18), and if necessary, changes have to be made to variance set $\{\nu_{i,j}^2\}_{j\in\mathcal{N}_i}$ in order to ensure the existence of the stabilizing solution to MARE (3.14). Then the optimal gain K_i can be obtained following (3.15b) or (3.25b), and consequently the MS stability holds for (3.18). Next we repeat the above steps for all $i \in \mathcal{N}$. Finally a sufficiently small $\varepsilon > 0$ can be chosen below the corresponding upper bound given in Theorem 6 or Corollary 1 and applied to (3.25) to complete the design of KCF for the entire sensor network.

3.4 Simulations

We demonstrate the performance of stationary DKF (3.15) and KCF (3.25) in the presence of data packet drops with two simulation examples.

Example 3.1. Consider a system of dynamics (3.12a) with system matrix A and initial state x_0 specified by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \overline{x}_0 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \Sigma_0 = 10 I_3.$$

Note that A has three different unstable eigenvalues on the unit circle, i.e., 1, $\pm j$. A sensor network of $N = 10$ sensor nodes, represented by the undirected topology in Figure 3.3, is employed for dynamic state estimation. The observation matrix for each sensor node is specified by

$$C_i = \begin{cases} C_o = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, & i \text{ is odd}, \\ C_e = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & i \text{ is even}, \end{cases}$$

for $i \in \mathcal{N}$. It is important to observe that the target state is undetectable by individual sensor nodes, however each inclusive neighbor set $\mathcal{J}_i$, $i \in \mathcal{N}$ contains nodes with observation matrices C_o and C_e, which guarantees the detectability of $(C_{\mathcal{J}_i}, A)$ $\forall$ $i \in \mathcal{N}$. The noise

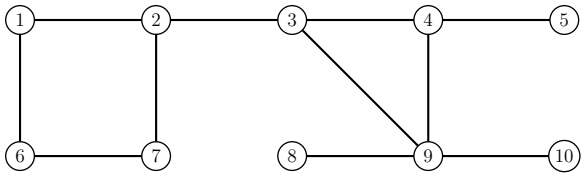


Figure 3.3: Network topology with $N = 10$ nodes

statistics are

$$Q = 0.005625 I_3, \;\; R_i = \begin{cases} 25\sqrt{i}, & i \text{ is odd}, \\ 25\sqrt{i} I_2, & i \text{ is even}. \end{cases}$$

The factor of $\sqrt{i}$ in the above expression allows individual nodes to have diverse noise conditions.

The data arrival rates for different communication channels between the neighboring nodes are selected as follows:

$$p_{1,2} = p_{2,1} = 0.3, \;\; p_{1,6} = p_{6,1} = 0.2, \;\; p_{2,3} = p_{3,2} = 0.25, \;\; p_{2,7} = p_{7,2} = 0.15,$$
$$p_{3,4} = p_{4,3} = 0.25, \;\; p_{3,9} = p_{9,3} = 0.1, \;\; p_{4,5} = p_{5,4} = 0.35, \;\; p_{4,9} = p_{9,4} = 0.1,$$
$$p_{6,7} = p_{7,6} = 0.35, \;\; p_{8,9} = p_{9,8} = 0.4, \;\; p_{9,10} = p_{10,9} = 0.4.$$

Note that $\gamma_{i,j}(k)$ and $\gamma_{j,i}(k)$ are independent of each other $\forall\, i \neq j$ despite that $p_{i,j} = p_{j,i}$ is adopted in our simulation setup. It is easy to verify by solving the LMI problem in Theorem 5 that, given the above statistics, the estimation error dynamics (3.18) are MS stabilizable, and thus MARE (3.14) admits a unique stabilizing solution $\forall\, i \in \mathcal{N}$. Initializing MDRE (3.17) with $\widehat{\Sigma}_i(0) = 0$ yields the limit that is the stabilizing solution to MARE (3.14). Then the optimal estimation gain K_i can be computed following (3.15b) or (3.25b) for both the stationary DKF and the KCF.

Following [16], we use the mean square deviation (MSD) to evaluate the performance of the algorithms. Let the MSD for node i at time index k be defined as

$$\mathrm{MSD}_i(k) := \mathrm{E}\left\{ \|x(k) - \widehat{x}_i(k)\|^2 \right\}.$$

Then the MSD for the entire network is defined as the average MSD over all nodes, i.e.,

$$\mathrm{MSD}(k) := \frac{1}{N}\sum_{i=1}^{N}\mathrm{MSD}_i(k).$$

In addition, the measurement [54]

$$\mathrm{diff}(k) := \sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\widehat{x}_i(k) - \overline{\widehat{x}}(k)\|^2}$$

is adopted to evaluate the disagreement among different local estimates, where

$$\overline{\widehat{x}}(k) := \frac{1}{N}\sum_{i=1}^{N}\widehat{x}_i(k)$$

is defined as the average state estimate over all nodes. In the simulations, the initial estimate is set as the mean value of the initial state by convention, i.e., $\widehat{x}_i(0) = \overline{x}_0 \ \forall \ i \in \mathcal{N}$. The simulation results are averaged over 1000 independent trials.

The transient network MSD performance of different algorithms over time index k is demonstrated in Figure 3.4. The algorithm marked by "SDKF" corresponds to the stationary DKF in (3.15), while the one denoted by "SDKF (No Packet Drops)" corresponds to a conventional stationary DKF where each node has access to the observations of its neighbors and no data packets are dropped during transmission. Similarly the algorithm marked by "KCF" represents KCF (3.25) with $\varepsilon = 0.25$, while the one denoted by "KCF (No Packet Drops)" is implemented by following the derivation of KCF (3.25) without including the presence of data packet drops [55]. The two algorithms named with "No Packet Drops" are included for comparison, which can help to evaluate the effect of frequent data drops on estimation accuracy. It can be observed from the plots that KCF outperforms SDKF by about 2 dB in this example, thus validating the MSD improvement introduced by the consensus term. In addition, KCF performs better than SDKF (No Packet Drops) as time proceeds, which modestly implies that the use of consensus term compensates the disadvantage posed

by the occurrence of data packet drops to some extent. Overall, KCF (No Packet Drops) possesses the best MSD performance since no information is lost and the consensus term is adopted.

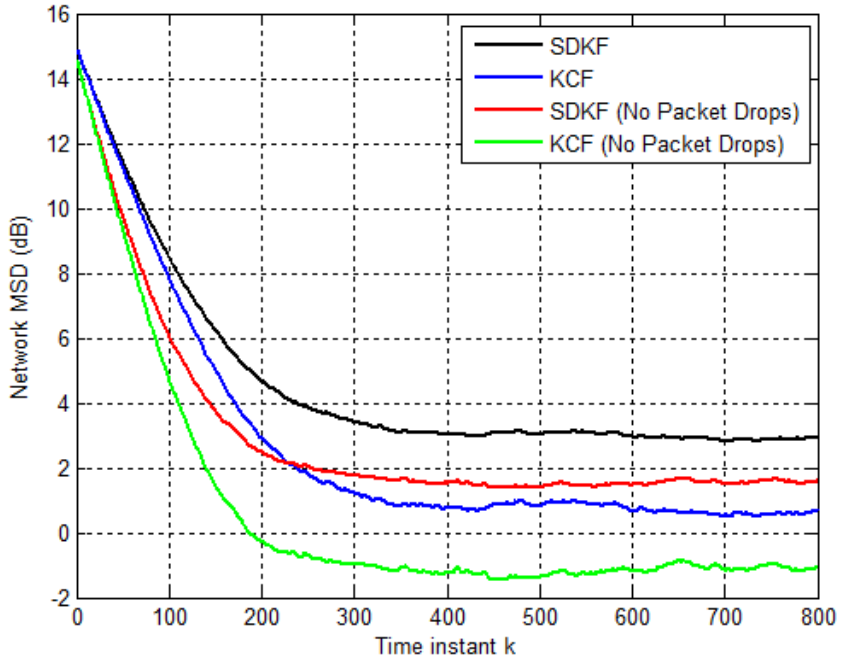


Figure 3.4: Example 3.1: MSD performance of different algorithms

Figure 3.5 shows the transient network MSD performance of KCF (3.25) when $\varepsilon = 0.02$, $\varepsilon = 0.25$, $\varepsilon = 0.5$, and $\varepsilon = 0.75$ are set respectively. Figure 3.6 compares the disagreement of local estimates with respect to the above different consensus coefficients. It is revealed from the plots that KCFs (3.25) with $\varepsilon = 0.25$ and $\varepsilon = 0.5$ have better MSD performance than those with $\varepsilon = 0.02$ and $\varepsilon = 0.75$, which suggests that MSD is not a monotonic function of the consensus coefficient. Another important feature is that although KCF with $\varepsilon = 0.25$ and that with $\varepsilon = 0.5$ give very close (almost the same) MSD performance, the latter embraces a smaller disagreement of local estimates than the former, which leads to a set of local estimates with higher cohesiveness. This result also corresponds to the aforementioned comment that although the consensus term may help to improve estimation accuracy, it is primarily used to achieve the consensus of estimates from different sensor nodes. A modest

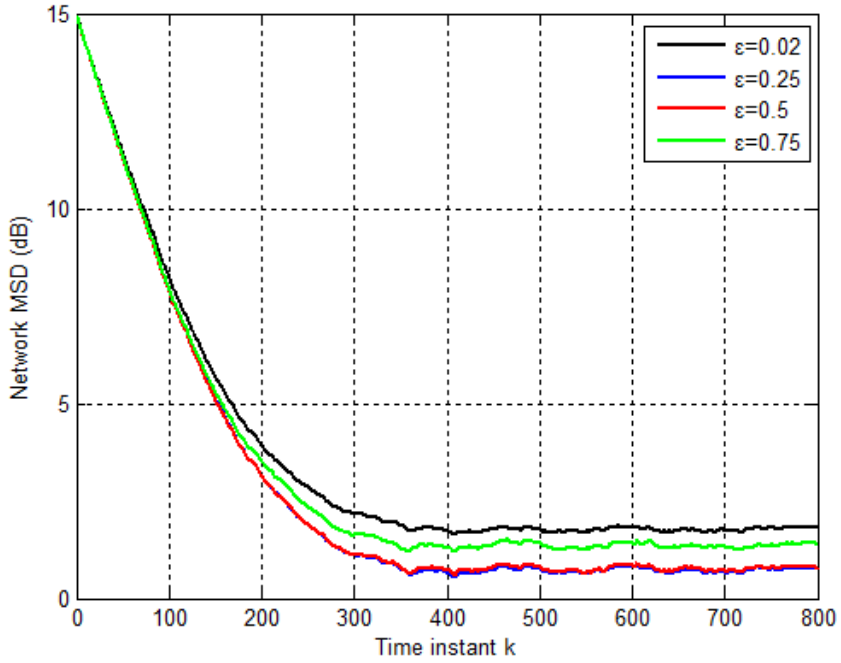


Figure 3.5: Example 3.1: MSD performance of KCF with different ε

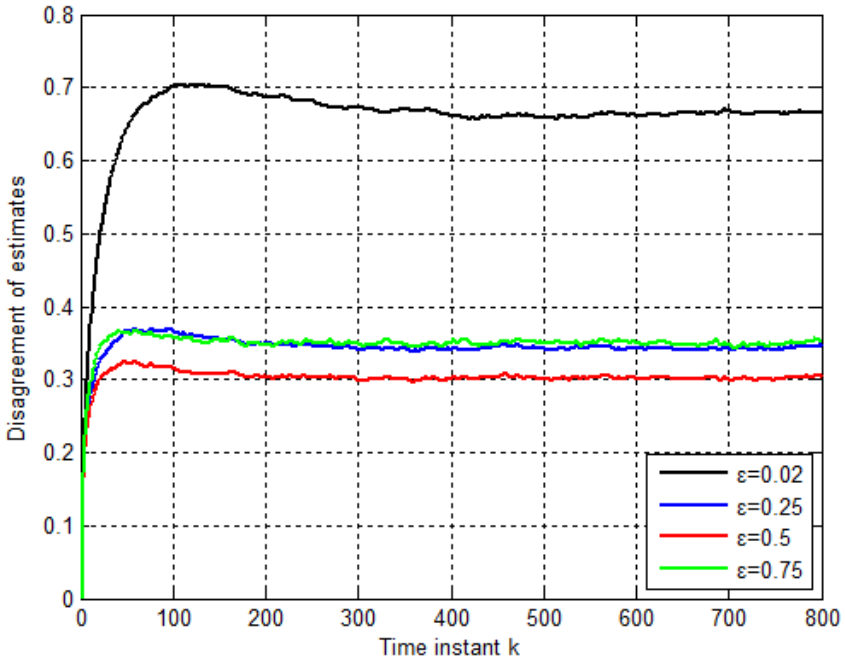


Figure 3.6: Example 3.1: Disagreement of local estimates under different ε

conclusion can be drawn from the above analysis that the consensus coefficient should be chosen carefully as a tradeoff between estimation accuracy and cohesiveness, and stability.

Example 3.2. The second example is adapted from the example in [16]. We consider the problem of estimating and tracking the position of a projectile. The position, velocity, and acceleration of the projectile are specified by

$$d = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

respectively, with the subscripts 1, 2 and 3 corresponding to the three spatial dimensions, and 3 being the vertical one. Hence the motion of the projectile is described by

$$v = \dot{d}, \quad a = \dot{v}, \quad a_1 = a_2 = 0, \quad a_3 = -g,$$

where g is the gravity constant. By taking

$$x_c = \begin{bmatrix} v \\ d \end{bmatrix}, \quad A_c = \begin{bmatrix} 0 & 0 \\ I_3 & 0 \end{bmatrix}, \quad b_c = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} \\ 0 \end{bmatrix},$$

the above can be written as the following continuous-time state space model

$$\dot{x}_c = A_c x_c + b_c.$$

Transforming the continuous-time model to discrete-time model with time-step T_s yields

$$x_c(t + T_s) = e^{T_s A_c} x_c(t) + \int_t^{t+T_s} e^{(t+T_s-\tau)A_c} b_c d\tau = (I + T_s A_c) x_c(t) + (T_s I + T_s^2 A_c/2) b_c.$$

Notice that the constant input $(T_s I + T_s^2 A_c/2) b_c$ can always be added to the filters to offset its impact on the state dynamics. Since we care about the estimation error instead of

the real state, this constant term can be disregarded for simulation convenience. Letting $x(k) = x_c(kT_s)$, $T_s = 0.1$, $A = I + T_s A_c$, and taking into account the effect of process noise, we arrive at the discrete-time state dynamics (3.12a). Note that A has six unstable eigenvalues on the unit circle.

In this example, a sensor network of $N = 20$ sensor nodes, represented by the undirected topology in Figure 3.7, is employed to estimate and track the position of the projectile. The observation matrix for each sensor node is specified by

$$C_i = \begin{cases} C_o = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, & i \text{ is odd}, \\ C_e = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, & i \text{ is even}, \end{cases}$$

for $i \in \mathcal{N}$. It is important to note that odd nodes can only measure the position of the moving target in the two horizontal dimensions, while even nodes can only observe the position in a combination of one horizontal dimension and the vertical dimension. Therefore the target is undetectable by individual sensor nodes. Nevertheless there is at least one node of each type in each inclusive neighbor set $\mathcal{J}_i$, $i \in \mathcal{N}$, so that the detectability of $(C_{\mathcal{J}_i}, A)$ is guaranteed for all $i \in \mathcal{N}$. The initial conditions and the noise statistics are given by

$$\overline{x}_0 = [10\ 2\ 8\ 0.1\ 0.1\ 0.1]', \quad \Sigma_0 = I_6, \quad Q = 0.001 I_6, \quad R_i = \begin{cases} \sqrt{i} \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}, & i \text{ is odd}, \\ \sqrt{i} \begin{bmatrix} 0.5 & 0 \\ 0 & 3.5 \end{bmatrix}, & i \text{ is even}. \end{cases}$$

The factor of $\sqrt{i}$ in the above expression allows individual nodes to have diverse noise conditions.

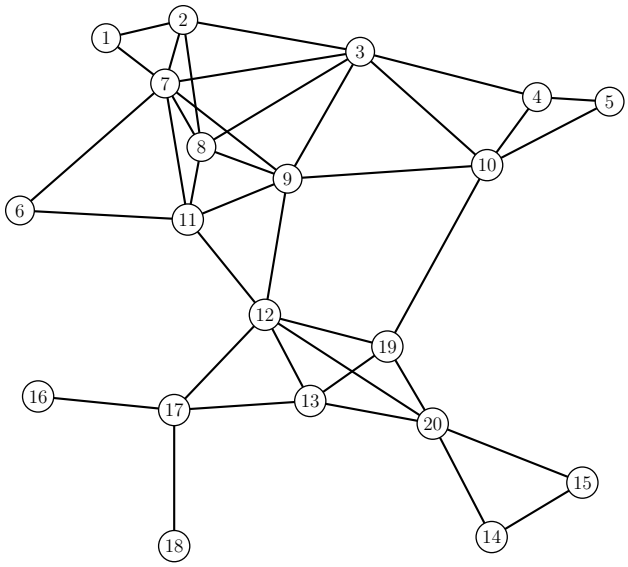


Figure 3.7: Network topology with $N = 20$ nodes

By solving the LMI problem in Theorem 5, we know that MS stabilizability holds for error dynamics (3.18) under the following data arrival rate setup for $i \in \mathcal{N}$, $j \in \mathcal{N}_i$:

$$p_{i,j} = \begin{cases} 0.6, & i \text{ is odd, } j \text{ is even OR } i \text{ is even, } j \text{ is odd,} \\ 0.2, & i \text{ is odd, } j \text{ is odd OR } i \text{ is even, } j \text{ is even.} \end{cases}$$

Thus MARE (3.14) admits a unique stabilizing solution, and the optimal gain K_i can be computed for both the stationary DKF and the KCF. Again, in the simulations, we set the initial estimate as the mean value of the initial state. The simulation results are averaged over 1000 independent trials.

Similar to Example 3.1, the transient network MSD performance of four different algorithms are compared. The simulation results shown in Figure 3.8 indicate that the proposed KCF improves over the proposed SDKF by about 3 dB and it even outperforms SDKF (No Packet Drops) in the steady-state by about 1.5 dB due to the presence of consensus term.

Figure 3.9 shows the transient network MSD performance of KCF (3.25) when $\varepsilon = 0.02$, $\varepsilon = 0.15$, $\varepsilon = 0.3$, and $\varepsilon = 0.45$ are set respectively. The disagreement of local estimates with respect to these consensus coefficients are depicted in Figure 3.10. It appears that KCF (3.25)

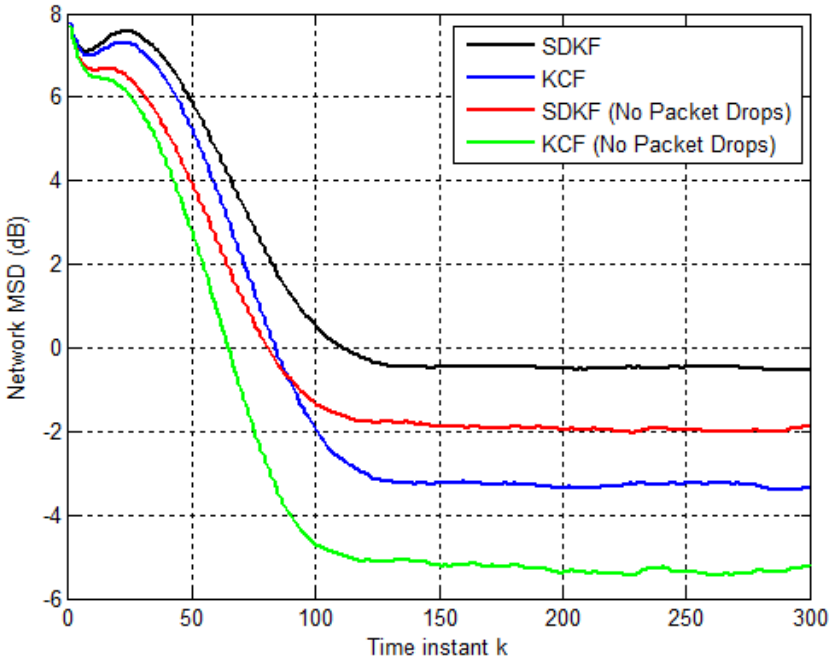


Figure 3.8: Example 3.2: MSD performance of different algorithms

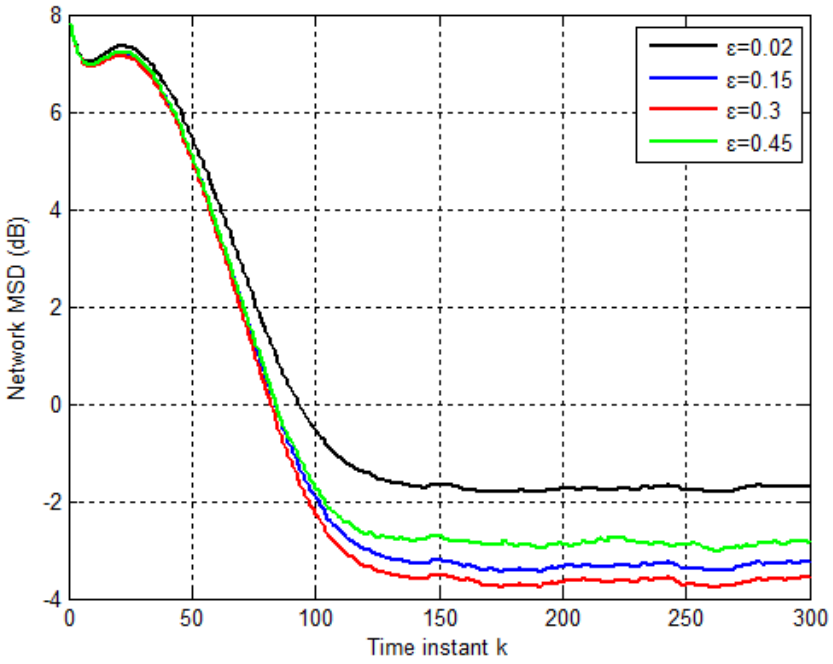


Figure 3.9: Example 3.2: MSD performance of KCF with different ε

with $\varepsilon = 0.3$ outperforms those with $\varepsilon = 0.02$, $\varepsilon = 0.15$, and $\varepsilon = 0.45$ since it produces not only the smallest network MSD in the steady-state but also a set of local estimates with the

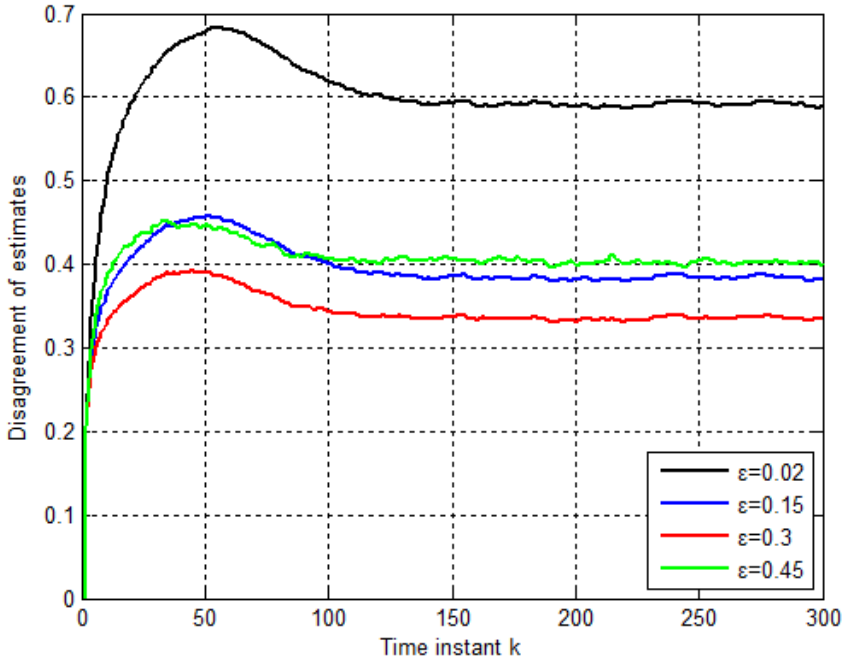


Figure 3.10: Example 3.2: Disagreement of local estimates under different ε

highest cohesiveness. The results from the plots also agree with the understanding that on one hand, the estimation performance may not be improved remarkably for small consensus coefficients (i.e., $\varepsilon = 0.02$); on the other hand, KCF tends to be closer to the unstable region with large consensus coefficients (i.e., $\varepsilon = 0.45$). These two situations should be balanced when we select the consensus coefficient for the KCF.

CHAPTER 4
DETERMINISTIC TIME-VARYING TOPOLOGIES

In this chapter we study state consensus control for discrete-time homogeneous MASs, and distributed state estimation over WSNs. Deterministic time-varying communication topologies are assumed in both problems.

4.1 Introduction

4.1.1 Elements of Graph Theory

Recall that in Chapter 3, the notions of neighbor set $\mathcal{N}_i$, stochastic time-varying graph $\mathcal{G}(k)$, Laplacian matrix $\mathcal{L}_\gamma(k)$, and more are adopted to address the issues of distributed Kalman filtering in the presence of data packet drops. All these notions are actually adapted from the classic graph theory. We start this chapter by reviewing some important terminologies and facts in graph theory in a detailed way, which will be employed to derive the main results in later sections. We use the following references [10, 12, 11] for the introduction. Other useful resources for graph theory will be cited as we proceed.

The communication between a network of interconnected systems (i.e., sensor nodes, autonomous agents, etc.) can be encoded through a weighted graph. The graph can be defined as directed or undirected. We focus on the introduction of directed graph (digraph), which can cover undirected graph as a special case in general, in spite of some noticeable difference. Let a weighted digraph be specified by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{v_i\}_{i=1}^N$ denoting the node set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ denoting the edge (arc) set, where an edge starting at node j (the parent node) and ending at node i (the child node) is denoted by $(v_j, v_i) \in \mathcal{E}$. Let $\mathcal{N} := \{1, \ldots, N\}$ denote the node index set. Then $\mathcal{N}_i = \{j \in \mathcal{N} | (v_j, v_i) \in \mathcal{E}\}$ denotes the set of neighbors of node i, and $\deg_i = |\mathcal{N}_i|$ stands for the degree (number of neighbors) of node i, with $|\cdot|$ representing the number of elements in a set. A path on the digraph is an ordered set of distinct nodes $\{v_{i_1}, \ldots, v_{i_\kappa}\}$ such that $(v_{i_{j-1}}, v_{i_j}) \in \mathcal{E}$ for $2 \leq j \leq \kappa$. If there is a path

from v_i to v_j, then v_j is said to be reachable from v_i, which is denoted as $v_i \to v_j$. A digraph $\mathcal{G}$ is called connected if there exists a node v_i in $\mathcal{G}$ such that $v_i \to v_j \ \forall \ j \neq i,\ j \in \mathcal{N}$, and v_i is called a connected node in $\mathcal{G}$. All the notions and notations mentioned above are also applicable to undirected graphs. However it should be noted that, different from digraphs, each edge in an undirected graph is an unordered pair of distinct nodes, i.e., edge (v_i, v_j) and edge (v_j, v_i) represents the same edge in an undirected graph $\mathcal{G}$. Therefore we can treat an undirected graph $\mathcal{G}$ having edge (v_i, v_j) equivalently as a digraph $\mathcal{G}$ having both edge (v_i, v_j) and edge (v_j, v_i). Then it is easy to understand that each node in a connected undirected graph is a connected node.

Let $\mathcal{A} = [a_{i,j}] \in \mathbb{R}^{N\times N}$ be the weighted adjacency matrix associated with weighted digraph or undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $a_{i,j} > 0$ represents the coupling strength of edge $(v_j, v_i) \in \mathcal{E}$, while $a_{i,j} = 0$ if edge $(v_j, v_i) \notin \mathcal{E}$. Self edges do not exist in our discussion, i.e., $a_{i,i} = 0 \ \forall \ i \in \mathcal{N}$. Then the Laplacian matrix $\mathcal{L} = [l_{i,j}] \in \mathbb{R}^{N\times N}$ associated with graph $\mathcal{G}$ is defined as

$$l_{i,j} = \begin{cases} \sum\limits_{k=1}^{N} a_{i,k}, & j = i, \\ -a_{i,j}, & j \neq i. \end{cases}$$

It is easy to observe that $\mathcal{L}\mathbf{1}_N = 0$, and thus $\mathcal{L}$ has at least one zero eigenvalue. In addition, all the eigenvalues of the Laplacian matrix lie on the closed right half plane, i.e., $\mathrm{Re}\{\lambda_i(\mathcal{L})\} \geq 0 \ \forall \ i \in \mathcal{N}$. In fact the only eigenvalues of $\mathcal{L}$ on the imaginary axis are zero in light of the well-known Greshgorin Circle Theorem [36]. For undirected graph $\mathcal{G}$, $\mathcal{A}$ and $\mathcal{L}$ are both symmetric matrices since $a_{i,j} = a_{j,i} \ \forall \ i, j \in \mathcal{N}$, and thus $\mathcal{L} \geq 0$. Other properties of Laplacian matrix can be found in [45, 56, 18]. The following fact holds for both digraphs and undirected graphs.

Lemma 11. *Let $\mathcal{L}$ be the Laplacian matrix associated with graph $\mathcal{G}$. Then $\mathcal{G}$ is connected if and only if 0 is a simple eigenvalue of $\mathcal{L}$.*

4.1.2 Literature Survey

This subsection summarizes the main results in Xiao et al. (2005) [92] and Jadbabaie et al. (2003) [38], which motivate the work in later sections.

- Xiao et al. (2005)

In [92] the authors study the estimation of an unknown (but constant) vector parameter $\theta \in \mathbb{R}^n$ using a network of N local sensor nodes. The noisy measurement $y_i \in \mathbb{R}^{m_i}$ obtained by node i is specified by

$$y_i = A_i\theta + v_i, \;\; i \in \mathcal{N}, \tag{4.1}$$

where $A_i \in \mathbb{R}^{m_i \times n}$ is a known observation matrix, different for each $i \in \mathcal{N}$ in general. The measurement noise v_i is a random variable with mean zero and covariance Σ_i. By convention it is assumed that $\{v_i\}_{i=1}^N$ are independent and jointly Gaussian distributed. Denote

$$y = \text{vec}\{y_1, \ldots, y_N\}, \;\; v = \text{vec}\{v_1, \ldots, v_N\}, \;\; A = \text{col}\{A_1, \ldots, A_N\}, \;\; \Sigma = \text{diag}(\Sigma_1, \ldots, \Sigma_N).$$

Then the collective measurements of all sensor nodes can be written as

$$y = A\theta + v.$$

Assume that $\sum_{i=1}^{N} m_i \geq n$ and $\text{rank}\{A\} = n$. The maximum-likelihood (ML) estimate of θ given measurement y is specified by

$$\hat{\theta}_{\text{ML}} = \left(A'\Sigma^{-1}A\right)^{-1} A'\Sigma^{-1}y = \left(\sum_{i=1}^{N} A_i'\Sigma_i^{-1}A_i\right)^{-1} \sum_{i=1}^{N} A_i'\Sigma_i^{-1}y_i, \tag{4.2}$$

which is unbiased, i.e., $\text{E}\{\hat{\theta}_{\text{ML}}\} = \theta$, with error covariance

$$\text{E}\left\{(\hat{\theta}_{\text{ML}} - \theta)(\hat{\theta}_{\text{ML}} - \theta)'\right\} = \left(A'\Sigma^{-1}A\right)^{-1}.$$

Note that if measurement noises $\{v_i\}_{i=1}^N$ are not jointly Gaussian distributed, (4.2) yields the linear minimum variance unbiased estimate of θ.

The distributed sensor fusion scheme is studied, where no data fusion center exists and no sensor nodes have the global knowledge of the time-varying network topology. Every node is required to ultimately obtain a good estimate of θ, i.e., $\hat{\theta}_{\mathrm{ML}}$, only by exchanging information with its neighbors and carrying out local computation. This scenario is quite useful since in many practical cases, sensor nodes are asked to execute multiple tasks and need an accurate estimate of the unknown parameter θ to make on-site decisions.

The difficulty of this problem also lies heavily on the fact that the communication topology is time-varying and may be disconnected most of the time due to the mobility or power constraints of the sensor network. Let the deterministic time-varying undirected graph representing the sensor network be specified by $\mathcal{G}(k) = \{\mathcal{V}, \mathcal{E}(k)\}$, where $\mathcal{E}(k)$ is the set of active edges at time index k. Then the neighbor set of node i at time k can be denoted by $\mathcal{N}_i(k) = \{j \in \mathcal{N} | (v_j, v_i) \in \mathcal{E}(k)\}$, and the degree of nodes i at time k can be denoted by $\deg_i(k) = |\mathcal{N}_i(k)|$. Since the sensor network consists of N nodes, there are only a finite number of possible graphs, say a total of φ, that $\mathcal{G}(k)$ can take. Denote the collection of all possible graphs by $\{\mathcal{G}_i\}_{i=1}^{\varphi}$, where $\mathcal{G}_i = (\mathcal{V}, \mathcal{E}_i)$. Their union is defined as a graph with node set $\mathcal{V}$ and edge set that is the union of $\{\mathcal{E}_i\}_{i=1}^{\varphi}$, i.e., $\cup_{i=1}^{\varphi}\mathcal{G}_i = (\mathcal{V}, \cup_{i=1}^{\varphi}\mathcal{E}_i)$, where "$\cup$" denotes the union operation. The graphs in $\{\mathcal{G}_i\}_{i=1}^{\varphi}$ are called jointly connected if their union $\cup_{i=1}^{\varphi}\mathcal{G}_i$ is a connected graph [38].

A distributed iterative scheme, based on distributed average consensus in the network, is first proposed and applied to a simplified case of (4.1) as

$$y_i = \theta + v_i, \quad i \in \mathcal{N},$$

where θ is an unknown (but constant) scalar parameter to be estimated, and the measurement noises $\{v_i\}_{i=1}^N$ are i.i.d. Gaussian variables with mean zero and variance σ^2. In this case the

ML estimate of θ happens to be the average of measurements obtained at all nodes, i.e.,

$$\hat{\theta}_{\mathrm{ML}} = \frac{1}{N}\mathbf{1}_N' y.$$

Therefore, in order to obtain a local estimate of θ, each node have to compute the average of all measurements $\{y_i\}_{i=1}^N$. By the distributed linear iterative scheme, each node initializes its local state $x_i(k)$ as $x_i(0) = y_i$ at time index $k = 0$, and then iteratively updates its state following the distributed protocol

$$x_i(k+1) = w_{i,i}(k)x_i(k) + \sum_{j\in\mathcal{N}_i(k)} w_{i,j}(k)x_j(k), \quad i \in \mathcal{N}, \quad k \geq 0. \tag{4.3}$$

Obviously the updated state at node i is a linear combination of its own prior state and the prior states at its instantaneous neighbors. Let $W(k) = [w_{i,j}(k)] \in \mathbb{R}^{N\times N}$ with $w_{i,j}(k) > 0$ for $j = i$, $j \in \mathcal{N}_i(k)$, and $w_{i,j}(k) = 0$ otherwise. By taking $x(k) = \mathrm{vec}\{x_1(k), \ldots, x_N(k)\}$, distributed protocol (4.3) can be written in a compact form as

$$x(k+1) = W(k)x(k) \tag{4.4}$$

with initial condition $x(0) = y$. Now the goal is to design weight matrix $W(k)$ following the sparsity pattern specified by the communication topology $\mathcal{G}(k)$, such that the states at all the nodes reach average consensus, i.e.,

$$\lim_{k\to\infty} x(k) = \hat{\theta}_{\mathrm{ML}}\mathbf{1}_N = \left(\frac{1}{N}\mathbf{1}_N'x(0)\right)\mathbf{1}_N. \tag{4.5}$$

Two simple rules for weight selection are proposed in [92], both of which lead to the average consensus (4.5), provided that the infinitely occurring communication graphs are jointly connected. The main result is given below.

Theorem 7. *Assume that the collection of communication graphs that occur infinitely often are jointly connected. Then iteration protocol* (4.4) *achieves convergence* (4.5) *for all* $x(0) \in \mathbb{R}^N$ *with either* (i) *the maximum-degree weights, specified by*

$$w_{i,j}(k) = \begin{cases} \frac{1}{N}, & j \in \mathcal{N}_i(k), \\ 1 - \frac{\deg_i(k)}{N}, & j = i, \\ 0, & \text{otherwise}, \end{cases}$$

or (ii) *the Metropolis weights, specified by*

$$w_{i,j}(k) = \begin{cases} \frac{1}{1+\max\{\deg_i(k),\, \deg_j(k)\}}, & j \in \mathcal{N}_i(k), \\ 1 - \sum\limits_{l \in \mathcal{N}_i(k)} w_{i,l}(k), & j = i, \\ 0, & \text{otherwise}. \end{cases}$$

The scheme for distributed sensor fusion is then proposed, based on average consensus, to tackle the ML estimation problem with the general setup in (4.1). In the scheme, each node initializes its local composite information matrix $P_i(k) \in \mathbb{R}^{n \times n}$ and local composite information state $q_i(k) \in \mathbb{R}^n$ as $P_i(0) = A_i' \Sigma_i^{-1} A_i$ and $q_i(0) = A_i' \Sigma_i^{-1} y_i$ at time index $k = 0$, respectively. Then the average consensus can be conducted entry-wise for $P_i(0)$ and $q_i(0)$ as

$$P_i(k+1) = w_{i,i}(k) P_i(k) + \sum_{j \in \mathcal{N}_i(k)} w_{i,j}(k) P_j(k), \;\; q_i(k+1) = w_{i,i}(k) q_i(k) + \sum_{j \in \mathcal{N}_i(k)} w_{i,j}(k) q_j(k),$$

respectively. Either the maximum-degree rule or the Metropolis rule can be adopted for weight selection. In light of Theorem 7, provided the joint connectedness of the infinitely occurring communication graphs, there hold

$$\lim_{k \to \infty} P_i(k) = \frac{1}{N} \sum_{j=1}^{N} A_j' \Sigma_j^{-1} A_j, \;\; \lim_{k \to \infty} q_i(k) = \frac{1}{N} \sum_{j=1}^{N} A_j' \Sigma_j^{-1} y_j.$$

As a result, sensor node i can compute the ML estimate in (4.2) asymptotically via

$$\hat{\theta}_{\mathrm{ML}} = \lim_{k\to\infty} P_i(k)^{-1} q_i(k), \quad i \in \mathcal{N}.$$

- Jadbabaie et al. (2003)

The consensus for MASs over time-varying communication topology has been widely studied [38, 57, 49, 67, 69, 50, 68, 99]. Here we introduce the main results in Jadbabaie et al. (2003) [38]. The authors study the coordination of a group of mobile autonomous agents, focusing on the state consensus for discrete-time homogeneous MASs over time-varying directed interaction graph $\mathcal{G}(k)$. The MAS consists of N dynamic agents described by

$$x_i(k+1) = x_i(k) + u_i(k), \quad x_i(0) = x_{i0} \tag{4.6}$$

for the ith agent, where $x_i(k) \in \mathbb{R}^n$ is the state vector, $u_i(k) \in \mathbb{R}^n$ is the control input, and $i \in \mathcal{N}$. We only present the analysis for the scalar case, i.e., $n = 1$, since it generalizes easily to the vector case by using the Kronecker product. Selecting the control protocol as

$$u_i(k) = \frac{1}{1 + \deg_i(k)} \sum_{j\in\mathcal{N}_i(k)} \left[x_j(k) - x_i(k)\right],$$

and substituting it into (4.6) yield

$$x_i(k+1) = \frac{1}{1 + \deg_i(k)} \left[x_i(k) + \sum_{j\in\mathcal{N}_i(k)} x_j(k)\right]. \tag{4.7}$$

Let $W(k) = [w_{i,j}(k)] \in \mathbb{R}^{N\times N}$ follow the *nearest neighbor rule*, specified by

$$w_{i,j}(k) = \begin{cases} \frac{1}{1+\deg_i(k)}, & j \in \mathcal{N}_i(k),\ j = i, \\ 0, & \text{otherwise.} \end{cases}$$

By taking $x(k) = \text{vec}\{x_1(k), \ldots, x_N(k)\}$, dynamic protocol (4.7) can be written in a compact form as

$$x(k+1) = W(k)x(k). \tag{4.8}$$

The main result is given below.

Theorem 8. *Assume that there exists an infinite sequence of contiguous, non-empty, bounded time intervals* $[k_j, k_{j+1})$, $j \in \{0, 1, \ldots\}$ *starting at* $k_0 = 0$ *such that* $\cup_{k=k_j}^{k_{j+1}-1} \mathcal{G}(k)$ *is connected for all* $j \in \{0, 1, \ldots\}$. *Then dynamic protocol* (4.8) *leads to state consensus with* $W(k)$ *following the nearest neighbor rule, and* $\lim_{k\to\infty} x(k) = c\mathbf{1}_N$ *for some consensus value* c.

Xiao et al. (2005) studies the distributed sensor fusion problem, while Jadbabaie et al. (2003) focuses on the consensus of MASs, both works end up investigating the consensus of states, governed by dynamic protocol (4.8), over deterministic time-varying communication graph. Xiao et al. (2005) assumes the graph to be undirected, and the proposed maximum-degree rule and Metropolis rule enable all the states to converge to the average of their initial values. Jadbabaie et al. (2003) considers time-varying digraph, and the state dynamics following the proposed nearest neighbor rule reach consensus at some value, but fail to achieve average consensus. In spite of the different results, matrix $W(k)$ is chosen to be paracontracting with respect to the Euclidean norm in both works, which proves to be a key point for consensus. A common drawback in Xiao et al. (2005) and Jadbabaie et al. (2003) is that the form of the proposed state dynamics is too restrictive, and thus the consensus protocols cannot be applied to systems of more general form. Generalized solutions to state consensus for discrete-time homogeneous MASs will be discussed in the next section.

4.2 State Consensus Control

In this section we investigate state consensus control for discrete-time homogeneous MASs over deterministic time-varying communication topology, which generalizes the work in Xiao et al. (2005) and Jadbabaie et al. (2003).

4.2.1 Problem Formulation and Preliminaries

The homogeneous MAS under our consideration consists of N discrete-time dynamic agents described by

$$x_i(k+1) = Ax_i(k) + Bu_i(k), \quad x_i(0) = x_{i0}, \tag{4.9}$$

for the ith agent where $x_i(k) \in \mathbb{R}^n$ is the state vector, $u_i(k) \in \mathbb{R}^m$ is the control input, and $i \in \mathcal{N}$. A challenging problem is state consensus over the time-varying feedback topology that may not be connected at any time $k \geq 0$. For leaderless state consensus, we consider the distributed state feedback control protocol

$$u_i(k) = -F \sum_{j=1}^{N} a_{i,j}(k)[x_i(k) - x_j(k)] \tag{4.10}$$

for $i \in \mathcal{N}$, assuming the accessability of the state vector and time-varying $a_{i,j}(k) \geq 0$ $\forall\, i, j \in \mathcal{N}$ as the entries of the adjacency matrix associated with the time-varying feedback graph $\mathcal{G}(k)$. Substituting (4.10) into (4.9) yields

$$x_i(k+1) = Ax_i(k) - BF \sum_{j=1}^{N} a_{i,j}(k)[x_i(k) - x_j(k)].$$

Let $x(k) = \text{vec}\{x_1(k), \ldots, x_N(k)\}$ be the global state vector. The feedback MAS admits the state space description

$$x(k+1) = [I_N \otimes A - \mathcal{L}(k) \otimes (BF)]x(k) \tag{4.11}$$

with $\mathcal{L}(k)$ being the Laplacian matrix associated with digraph $\mathcal{G}(k)$. The state consensus control requires design of the state feedback gain F that achieves

$$\lim_{k\to\infty} [x_i(k) - x_j(k)] = 0 \;\; \forall\, i, j \in \mathcal{N}. \tag{4.12}$$

The eigenvalues of $\mathcal{L}(k)$ can be arranged in ascending order according to their absolute values, i.e.,

$$0 = \lambda_1\{\mathcal{L}(k)\} \leq |\lambda_2\{\mathcal{L}(k)\}| \leq \cdots \leq |\lambda_N\{\mathcal{L}(k)\}|.$$

The average of the Laplacian matrix over time interval $[k,\ k+T)$ with integer $T > 0$ is defined by

$$\overline{\mathcal{L}}(k) := \frac{1}{T}\sum_{i=0}^{T-1} \mathcal{L}(k+i). \tag{4.13}$$

The graph $\overline{\mathcal{G}}(k)$ corresponding to the average Laplacian matrix $\overline{\mathcal{L}}(k)$ can be interpreted as the *union graph* over time interval $[k,\ k+T)$. The uniform connectedness is defined next in light of Lemma 11.

Definition 4. *A time-varying digraph is uniformly connected, if there exists a finite $T > 0$ such that the average Laplacian matrix defined in* (4.13) *satisfies $|\lambda_2\{\overline{\mathcal{L}}(k)\}| > 0\ \forall\ k \geq 0$.*

For a time-varying digraph, its Laplace matrix $\mathcal{L}(k)$ is asymmetric and can have complex eigenvalues and multiple zero eigenvalues. Nevertheless there always exists one eigenvector $v_1 = \frac{\mathbf{1}_N}{\sqrt{N}}$ corresponding to one zero eigenvalue. Set $\{v_i\}_{i=2}^N$ such that $\{v_i\}_{i=1}^N$ form an orthonormal basis for $\mathbb{R}^N$. Denote $V = [\ v_1\ \ \cdots\ \ v_N\ ]$, $\widehat{V} = [\ v_2\ \ \cdots\ \ v_N\ ] \in \mathbb{R}^{N\times(N-1)}$. It follows that

$$V'\mathcal{L}(k)V = \begin{bmatrix} v_1'\mathcal{L}(k)v_1 & v_1'\mathcal{L}(k)\widehat{V} \\ \widehat{V}'\mathcal{L}(k)v_1 & \widehat{V}'\mathcal{L}(k)\widehat{V} \end{bmatrix} = \begin{bmatrix} 0 & \ell(k) \\ 0 & \widehat{\mathcal{L}}(k) \end{bmatrix}\ \ \forall\ k \geq 0,$$

where $\widehat{\mathcal{L}}(k) \in \mathbb{R}^{(N-1)\times(N-1)}$ is a reduced Laplacian matrix, and $\ell(k)' \in \mathbb{R}^{N-1}$. Recall the property of Kronecker product [9]:

$$(M_1 \otimes M_2)(M_3 \otimes M_4) = (M_1M_3) \otimes (M_2M_4)$$

whenever the dimensions make sense. Let $\widetilde{V} = V \otimes I_n$. Then

$$\widetilde{V}'[\mathcal{L}(k) \otimes BF]\widetilde{V} = (V' \otimes I_n)[\mathcal{L}(k) \otimes (BF)](V \otimes I_n) = [V'\mathcal{L}(k)V] \otimes (BF).$$

Define

$$\overline{x}(k) := (v_1' \otimes I_n)x(k) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_i(k) \in \mathbb{R}^n, \quad \widetilde{x}(k) := (\widehat{V}' \otimes I_n)x(k) \in \mathbb{R}^{(N-1)n}.$$

Applying similarity transform $\widetilde{V}'$ to the collective state space model in (4.11) yields

$$\overline{x}(k+1) = A\overline{x}(k) - [\ell(k) \otimes (BF)]\widetilde{x}(k), \tag{4.14}$$

$$\widetilde{x}(k+1) = [I_{N-1} \otimes A - \widehat{\mathcal{L}}(k) \otimes (BF)]\widetilde{x}(k). \tag{4.15}$$

The following is the main result of this subsection.

Theorem 9. *Feedback MAS* (4.11) *achieves the state consensus defined in* (4.12)*, if and only if the dynamic system described in* (4.15) *is asymptotically stable.*

Proof. Suppose that the dynamic system described in (4.15) is asymptotically stable. Then

$$\lim_{k\to\infty} \widetilde{x}(k) = 0, \quad \lim_{k\to\infty} \overline{x}(k) = \lim_{k\to\infty} A^{(k-k_0)}\overline{x}_0$$

for some vector $\overline{x}_0 \neq 0$ and integer $k_0 \geq 0$. That is, $\overline{x}(k)$ approaches the solution of $\overline{x}(k+1) = A\overline{x}(k)$ under initial condition $\overline{x}(k_0) = \overline{x}_0$. It follows that

$$\begin{aligned}\lim_{k\to\infty} x(k) &= \lim_{k\to\infty} \widetilde{V} \begin{bmatrix} \overline{x}(k) \\ \widetilde{x}(k) \end{bmatrix} = \lim_{k\to\infty} \widetilde{V} \begin{bmatrix} \overline{x}(k) \\ 0 \end{bmatrix} = \lim_{k\to\infty} (v_1 \otimes I_n)\overline{x}(k) \\ &= (v_1 \otimes I_n) \lim_{k\to\infty} A^{(k-k_0)}\overline{x}(k_0) = (\mathbf{1}_N \otimes I_n) \lim_{k\to\infty} A^{(k-k_0)} \left[\frac{1}{N}\sum_{i=1}^{N} x_i(k_0)\right].\end{aligned}$$

Hence $x_i(k) \to A^{(k-k_0)} \left[\frac{1}{N} \sum_{j=1}^{N} x_j(k_0) \right] \ \forall\ i \in \mathcal{N}$ as $k \to \infty$, which validates the state consensus defined in (4.12).

Conversely if state consensus is achieved, then the state vector of the dynamic system described in (4.11) satisfies

$$\lim_{k\to\infty} x(k+1) = \lim_{k\to\infty} (I_N \otimes A)x(k),$$

and $x_i(k) \to \frac{\overline{x}(k)}{\sqrt{N}}$ for each $i \in \mathcal{N}$ and some $\overline{x}(k)$ that is a solution to $\overline{x}(k+1) = A\overline{x}(k)$ under an appropriate initial condition. Denote $\tilde{x}(k) = x(k) - v_1 \otimes \overline{x}(k)$ as the global consensus error. Its dynamics are governed by

$$\begin{aligned}
\tilde{x}(k+1) &= [I_N \otimes A - \mathcal{L}(k) \otimes (BF)]x(k) - v_1 \otimes [A\overline{x}(k)] \\
&= [I_N \otimes A - \mathcal{L}(k) \otimes (BF)]x(k) - (I_N \otimes A)[v_1 \otimes \overline{x}(k)] + [\mathcal{L}(k)v_1] \otimes [BF\overline{x}(k)] \\
&= [I_N \otimes A - \mathcal{L}(k) \otimes (BF)]x(k) - [I_N \otimes A - \mathcal{L}(k) \otimes (BF)][v_1 \otimes \overline{x}(k)] \\
&= [I_N \otimes A - \mathcal{L}(k) \otimes (BF)]\, \tilde{x}(k)
\end{aligned}$$

that is the same as (4.11) except that $x(k)$ is replaced by $\tilde{x}(k)$. Applying the same similarity transform $\widetilde{V}'$ to the above state space model yields two similar state space equations to those of (4.14) and (4.15) respectively, with the latter given by

$$\widetilde{\tilde{x}}(k+1) = [I_{N-1} \otimes A - \widehat{\mathcal{L}}(k) \otimes (BF)]\widetilde{\tilde{x}}(k).$$

The above is the same as (4.15) except that $\widetilde{x}(k)$ is replaced by $\widetilde{\tilde{x}}(k)$. The hypothesis on the state consensus implies that the consensus error $\tilde{x}(k) \to 0$ and thus $\widetilde{\tilde{x}}(k) \to 0$ as $k \to \infty$. It follows that the dynamic system described in (4.15) is asymptotically stable, thereby concluding the proof. □

Denote $\widetilde{A} = I_{N-1} \otimes A$, $\widetilde{B} = I_{N-1} \otimes B$, $\widetilde{F} = I_{N-1} \otimes F$, and $\widehat{\mathcal{L}}_m(k) = \widehat{\mathcal{L}}(k) \otimes I_m$. Then the dynamic system described in (4.15) can be rewritten as

$$\widetilde{x}(k+1) = \left[\widetilde{A} - \widetilde{B}\widehat{\mathcal{L}}_m(k)\widetilde{F}\right]\widetilde{x}(k). \tag{4.16}$$

Similarly denote transfer matrices $G(z) = F(zI - A)^{-1}B$, and $\widetilde{G}(z) = I_{N-1} \otimes G(z)$. Then dynamic system (4.15) has the feedback form as configured in Figure 4.1. The state consensus is now equivalent to the asymptotic stability of the time-varying feedback system in Figure 4.1 in light of Theorem 9.

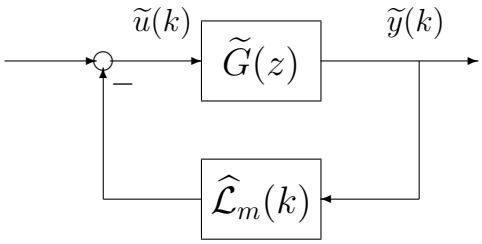


Figure 4.1: Time-varying dynamic system (4.15) in feedback form

The definition of semisimple eigenvalue is provided next.

Definition 5. [28] *An eigenvalue of matrix A, denoted by $\lambda_i(A)$, is said to be semisimple, if the algebraic multiplicity of $\lambda_i(A)$, which is its multiplicity as a root of the corresponding characteristic polynomial, equals the geometric multiplicity of $\lambda_i(A)$, which is the maximum number of linearly independent eigenvectors associated with $\lambda_i(A)$.*

It follows that a simple eigenvalue is always semisimple, but not conversely. For convenience, the next assumption is made throughout this section.

Assumption 6. *All eigenvalues of the system matrix A lie on the unit circle.*

Because stable eigenvalues do not require control actions, the assumption on the system matrix A has no loss of generality. Under Assumption 6, the MAS described in (4.9) is said to be neutrally stable if matrix A only has semisimple eigenvalues, and it is said to be neutrally unstable if A contains non-semisimple eigenvalues. The state consensus results for

neutrally stable MAS and neutrally unstable MAS will be given respectively in the following two subsections.

4.2.2 Neutrally Stable MAS

While Theorem 9 provides an equivalent state consensus condition, the challenge remains to design of the state feedback gain F that stabilizes dynamic system (4.15) over the time-varying graph. In this subsection we restrict feedback graphs to the following set of time-varying graphs.

Assumption 7. *The time-varying graph $\mathcal{G}(k)$ with Laplacian matrix $\mathcal{L}(k)$ is undirected, satisfying* (a) $0 \le \mathcal{L}(k) \le I \ \forall\ k \ge 0$, *and* (b) $\mathcal{G}(k)$ *is uniformly connected, i.e., there exists a finite $T > 0$ such that*

$$\overline{\widehat{\mathcal{L}}}_m(k) := \frac{1}{T}\sum_{i=0}^{T-1} \widehat{\mathcal{L}}_m(k+i) > 0 \ \ \forall\ k \ge 0.$$

For (a), it is noted that $\mathcal{L}(k) \ge 0$ follows from the property of the undirected graph, and $\mathcal{L}(k) \le I$ can be satisfied by appropriate (positive and symmetric) scaling. In addition, let "$\Rightarrow$" stand for "implies". There holds

$$0 \le \mathcal{L}(k) \le I_N \ \ \Rightarrow \ \ 0 \le \widehat{V}'\mathcal{L}(k)\widehat{V} \le I_N \ \ \Rightarrow \ \ 0 \le \widehat{\mathcal{L}}(k) \le I_{N-1} \ \ \Rightarrow \ \ 0 \le \widehat{\mathcal{L}}_m(k) \le I_{(N-1)m}.$$

For (b), the uniform connectedness of $\mathcal{G}(k)$ implies that

$$V'\overline{\mathcal{L}}(k)V = \frac{1}{T}\sum_{i=0}^{T-1} V'\mathcal{L}(k+i)V = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{T}\sum\limits_{i=0}^{T-1}\widehat{\mathcal{L}}(k+i) \end{bmatrix}$$

satisfies $\lambda_2\{V'\overline{\mathcal{L}}(k)V\} = \lambda_2\{\overline{\mathcal{L}}(k)\} > 0\ \forall\ k \ge 0$. It follows that

$$\overline{\widehat{\mathcal{L}}}(k) := \frac{1}{T}\sum_{i=0}^{T-1} \widehat{\mathcal{L}}(k+i) > 0 \ \ \forall\ k \ge 0,$$

which is equivalent to the inequality in (b). For the agent system under state feedback represented by $G(z) = F(zI - A)^{-1}B$, the following assumption is made:

Assumption 8. *The pair (A, B) is reachable, and there exists matrix $X > 0$ satisfying*

$$\text{(a) } X = A'XA, \quad \text{(b) } I - B'XB \geq 0, \quad \text{(c) } F = B'XA.$$

It is important to observe that the existence of X requires A only having semisimple eigenvalues. Moreover, in light of [91], there is no loss of generality to consider system matrix of form $A = \text{diag}(A_1, \cdots, A_\ell)$, where $A_i = 1$ or $A_i = -1$, or

$$A_i = \begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{bmatrix}, \quad \theta_i \in [0,\ 2\pi).$$

It follows from (a) that X is block diagonal with the ith diagonal block being $\rho_i > 0$ or $\rho_i I_2 > 0$, and $\{\rho_i\}_{i=1}^{\ell}$ can be chosen to satisfy (b). The feedback gain is highlighted by (c). For A of general form, similarity transformation can be applied first for the diagonalization of A. It is easy to verify that $A - BF$ is a Schur stability matrix.

Lemma 12. *Under Assumption 6 and 8, the transfer matrix $G(z) + \frac{1}{2}I = F(zI - A)^{-1}B + \frac{1}{2}I$ is positive real (PR), i.e.,*

$$\left[G(z) + \frac{1}{2}I\right]^* + \left[G(z) + \frac{1}{2}I\right] \geq 0 \ \ \forall\, |z| \geq 1,$$

and (F, A) is observable.

Proof. Under Assumption 6 and 8, there holds

$$\begin{aligned} &G(z)^* + G(z) + I \geq G(z)^* + G(z) + B'XB \\ &= B'[(zI - A)^*]^{-1}A'XB + B'XA(zI - A)^{-1}B + B'XB \\ &= B'[(zI - A)^*]^{-1}\left[A'X(zI - A) + (zI - A)^*XA + (zI - A)^*X(zI - A)\right](zI - A)^{-1}B \end{aligned}$$

$$= B'[(zI - A)^*]^{-1}\,(z^*zX - A'XA)\,(zI - A)^{-1}B$$
$$= \left(|z|^2 - 1\right) B'[(zI - A)^*]^{-1}X(zI - A)^{-1}B \geq 0 \quad \forall\, |z| \geq 1.$$

For those $z = \lambda_i(A)$, limit can be taken with respect to $z \to \lambda_i(A)$, and the above inequality still holds. We thus conclude that $G(z) + \frac{1}{2}I$ is PR.

To show the observability of (F, A), we let $X = U'U > 0$ for some nonsingular U. Under the similarity transform U, realization (A, B, F) is transformed to (UAU^{-1}, UB, FU^{-1}). The reachability of (A, B) is equivalent to that of (UAU^{-1}, UB), which leads to the observability of $\{(UB)', (UAU^{-1})'\}$. Equations (a) and (c) in Assumption 8 lead to

$$U'U = A'U'UA \;\Rightarrow\; (UAU^{-1})'(UAU^{-1}) = I, \quad F = B'U'UA \;\Rightarrow\; FU^{-1} = (UB)'(UAU^{-1}).$$

We thus have the equivalence of the observability for $\{(UB)', (UAU^{-1})'\}$ to that for

$$\{(UB)'(UAU^{-1}), (UAU^{-1})^{-1}(UAU^{-1})'(UAU^{-1})\} = (FU^{-1}, UA^{-1}U^{-1})$$

under the similarity transform $(UAU^{-1})^{-1}$. Then applying the similarity transform U^{-1} yields the observability for (F, A^{-1}), which is equivalent to that for (F, A) given the nonsingularity of A, thereby concluding the proof. □

Denote $\widehat{\Delta}(k) = I - \widehat{\mathcal{L}}_m(k)$. Assumption 7 implies that

$$0 \leq \widehat{\Delta}(k) \leq I \;\; \forall\, k \geq 0.$$

Hence the state equation in (4.16) can be rewritten as

$$\widetilde{x}(k+1) = \left[\left(\widetilde{A} - \widetilde{B}\widetilde{F}\right) + \widetilde{B}\widehat{\Delta}(k)\widetilde{F}\right]\widetilde{x}(k),$$

leading to the equivalent feedback system in Figure 4.2 where

$$\widetilde{T}_F(z) = \widetilde{F}\left(zI - \widetilde{A} + \widetilde{B}\widetilde{F}\right)^{-1}\widetilde{B}. \tag{4.17}$$

It is noted that $\widetilde{y}(k)$ in Figure 4.2 is the same as that in Figure 4.1, while $\widetilde{v}(k)$ in Figure 4.2 is related to $\widetilde{u}(k)$ in Figure 4.1 via $\widetilde{v}(k) = \widetilde{u}(k) + \widetilde{F}\widetilde{x}(k)$.

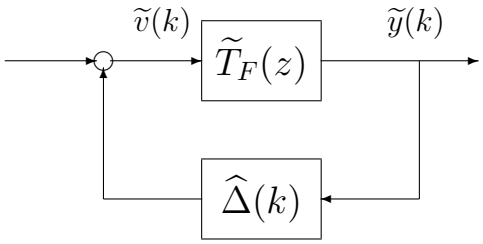


Figure 4.2: Equivalent feedback system to that in Figure 4.1

Let $\|T\|_{\mathcal{H}_\infty} := \sup_{|z|\geq 1} \overline{\sigma}[T(z)]$ be the $\mathcal{H}_\infty$ norm of a given transfer matrix $T(z)$. The next result shows that the transfer matrix $\widetilde{T}_F(z)$ in (4.17) is bounded real [5, 86].

Lemma 13. *Under Assumption* 6 *and* 8*, the transfer matrix* $\widetilde{T}_F(z)$ *in* (4.17) *satisfies* $\|\widetilde{T}_F\|_{\mathcal{H}_\infty} = 1$.

Proof. The expression of $\widetilde{T}_F(z)$ in (4.17) implies that

$$\widetilde{T}_F(z) = I_{N-1} \otimes T_F(z), \quad T_F(z) = F(zI - A + BF)^{-1}B.$$

Hence $\|\widetilde{T}_F\|_{\mathcal{H}_\infty} = \|T_F\|_{\mathcal{H}_\infty}$. It is now a known fact [32, 64] that

$$\inf_F \|T_F\|_{\mathcal{H}_\infty} = M(A) := \prod_{i=1}^{n} \max\{|\lambda_i(A)|, 1\} = 1,$$

where $M(A)$ is the Mahler measure of matrix A. To show that the infimum is achieved by those F satisfying Assumption 8, we note first that

$$T_F(z) = F(zI - A)^{-1}B[I + F(zI - A)^{-1}B]^{-1} = G(z)[I + G(z)]^{-1}.$$

We note next the equivalence of $\|T_F\|_{\mathcal{H}_\infty} = 1$ to

$$T_F(z)^* T_F(z) \leq I \ \ \forall \ |z| \geq 1. \tag{4.18}$$

The proof of the lemma can then be completed by noting the following equivalent relations:

$$\begin{aligned} (4.18) \quad &\Longleftrightarrow \quad G(z)^* G(z) \leq [I + G(z)]^* [I + G(z)] \ \ \forall \ |z| \geq 1 \\ &\Longleftrightarrow \quad \left[G(z) + \frac{1}{2} I\right]^* + \left[G(z) + \frac{1}{2} I\right] \geq 0 \ \ \forall \ |z| \geq 1. \end{aligned}$$

The last inequality implies the PR of $G(z) + \frac{1}{2} I$, which holds in light of Lemma 12. □

The next result is derived directly from the well-known Small Gain Theorem [27, 104].

Lemma 14. *Consider the closed-loop system configured in Figure* 4.2*. Assume that the transfer matrix* $\widetilde{T}_F(z)$ *is internally stable. Then the feedback system in Figure* 4.2 *is internally stable if* $\|\widetilde{T}_F\|_{\mathcal{H}_\infty} \|\widehat{\Delta}(k)\|_{\mathcal{H}_\infty} < 1$.

Recall that we require the time-varying graphs to be uniformly connected in Assumption 7. Nevertheless uniformly connected graphs are not adequate for achieving the state consensus. Let $\widehat{J}(k) \geq 0$ such that $\widehat{\mathcal{L}}_m(k) = \widehat{J}(k)^2$. The following assumption is also made.

Assumption 9. *The pair* $\{\widehat{J}(k)\widetilde{F}, \widetilde{A}\}$ *is uniformly observable in the sense that there exists a finite* $T_o > 0$ *such that*

$$\sum_{i=0}^{T_o - 1} \left[\widetilde{F} \widetilde{A}^i\right]' \widehat{\mathcal{L}}_m(k+i) \left[\widetilde{F} \widetilde{A}^i\right] > 0 \ \ \forall \ k \geq 0. \tag{4.19}$$

The state consensus result is presented below.

Theorem 10. *Under Assumptions* $6 \sim 9$*, the feedback homogeneous MAS described in* (4.11) *achieves the state consensus defined in* (4.12).

Proof. Because of the equivalence of the state consensus to the asymptotic stability of the feedback system in Figure 4.2, we will show that Assumptions $6 \sim 9$ ensure such a feedback

stability. For this purpose we first make use of the fact that Assumption 6 and 8 implies $\|\widetilde{T}_F\|_{\mathcal{H}_\infty} = 1$ in light of Lemma 13, and the fact that Assumption 9 implies the existence of a finite T_o such that inequality (4.19) holds. The proof can then be completed by showing that $\widehat{\Delta}(k)$ is a strict contraction map from $\ell_2^m[k_0,\ k_0+T_\mathrm{o}-1]$ to $\ell_2^m[k_0,\ k_0+T_\mathrm{o}-1]$ for all $k_0 \geq 0$, in light of Lemma 14. Since $\widehat{\Delta}(k)$ is symmetric satisfying $0 \leq \widehat{\Delta}(k) \leq I$ for all $k \geq 0$ by Assumption 7,

$$\widehat{\Delta}(k)^2 = \left[I - \widehat{\mathcal{L}}_m(k)\right]^2 = I - \widehat{\Gamma}(k) \geq 0 \ \ \forall\ k \geq 0,$$

where $\widehat{\Gamma}(k) = 2\widehat{\mathcal{L}}_m(k) - \widehat{\mathcal{L}}_m(k)^2$ satisfying

$$\widehat{\Gamma}(k) = \widehat{J}(k)\left[2I - \widehat{\mathcal{L}}_m(k)\right]\widehat{J}(k) \geq \widehat{J}(k)^2 = \widehat{\mathcal{L}}_m(k)$$

in light of Assumption 7. It follows that

$$\begin{aligned}
\sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \|\widetilde{v}(k)\|^2 &= \sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \widetilde{y}(k)'\widehat{\Delta}(k)^2\widetilde{y}(k) \\
&= \sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \widetilde{y}(k)'\left[I - \Gamma(k)\right]\widetilde{y}(k) \\
&\leq \sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \widetilde{y}(k)'\left[I - \widehat{\mathcal{L}}_m(k)\right]\widetilde{y}(k) \\
&= \sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \left[\|\widetilde{y}(k)\|^2 - \widetilde{y}(k)'\widehat{\mathcal{L}}_m(k)\widetilde{y}(k)\right] \\
&< \sum_{k=k_0}^{k_0+T_\mathrm{o}-1} \|\widetilde{y}(k)\|^2 \quad \forall\ k_0 \geq 0.
\end{aligned}$$

The last inequality follows from the uniform observability in (4.19). Indeed this inequality is violated only if

$$\widetilde{y}(k)'\widehat{\mathcal{L}}_m(k)\widetilde{y}(k) = 0 \ \ \forall\ k \in [k_0,\ k_0+T_\mathrm{o}-1],$$

where $\widetilde{y}(k) = \widetilde{F}\widetilde{A}^{(k-k_0)}\widetilde{x}_0$ according to Figure 4.1, with $\widetilde{x}_0 \neq 0$ denoting the initial condition at time index k_0, and hence it contradicts to the uniform observability. The above proves that $\widehat{\Delta}(k)$ is a strict contraction map from $\ell_2^m[k_0,\ k_0+T_{\rm o}-1]$ to $\ell_2^m[k_0,\ k_0+T_{\rm o}-1]$ for all $k_0 \geq 0$, thereby concluding the proof. □

Remark 5. While Theorem 10 establishes the state consensus, the convergence rate may not be exponential, which requires the exponential stability for the feedback system in Figure 4.2, in general. The reason lies in the fact that the limit of $\inf \|\widehat{\mathcal{L}}(k)\|$ can be zero. For instance $\mathcal{L}(k) = (1+k^2)^{-1}\mathcal{L}$ with $\mathcal{L}$ a constant Laplacian matrix corresponding a connected graph. In this case $\mathcal{L}(k) \to 0$ as $k \to \infty$, which cannot achieve exponential rate for the state consensus. If the exponential convergence rate is required, then the notion of the uniform observability in (4.19) needs to be strengthened to

$$\sum_{i=0}^{T_{\rm o}-1} \left[\widehat{F}\widehat{A}^i\right]' \widehat{\mathcal{L}}_m(k+i) \left[\widehat{F}\widehat{A}^i\right] \geq \epsilon_{\rm o} I \quad \forall\ k \geq 0$$

and for some $\epsilon_{\rm o} > 0$, which in turn requires $\lambda_2\{\overline{\mathcal{L}}(k)\} \geq \epsilon_{\rm c}\ \forall\ k \geq 0$ and for some $\epsilon_{\rm c} > 0$. In fact the above strengthened notion of uniform observability is adopted to address the state consensus control for neutrally unstable MAS. More details will be discussed in the next subsection. □

It is important to point out that Theorem 10 also solves the state consensus problem for $x(k+1) = W(k)x(k)$ over time-varying communication topology as discussed in Xiao et al. (2005) [92] and Jadbabaie et al. (2003) [38]. Specifically we can adopt $A = 1$, $B = 1$, $F = 1$, and thus the uniform observability in (4.19) reduces to the uniform connectedness in Assumption 7. Then $W(k) = I_N - \mathcal{L}(k)$ can be set with $\mathcal{L}(k)$ satisfying Assumption 7, which not only agrees with, but also generalizes the weight selection rules for $W(k)$ in [92]. It is shown in the proof of Theorem 9 that the average consensus required in [92] is also achieved, i,e., $x_i(k) \to \frac{1}{N}\sum_{j=1}^{N} x_j(0)\ \forall\ i \in \mathcal{N}$ as $k \to \infty$.

4.2.3 Neutrally Unstable MAS

Now we consider the case where A has at least one non-semisimple eigenvalue. Assumption 7 is imposed again while Assumption 8 no longer holds. The following lemma is useful.

Lemma 15. [31, 33] *Consider $T_F(z) = F(zI - A + BF)^{-1}B$ as the complementary sensitivity under state feedback where $F', B \in \mathbb{R}^{n\times m}$. If the pair (A, B) is stabilizable, then*

$$M(A)^{\frac{1}{m}} \leq \inf_F \|T_F\|_{\mathcal{H}_\infty} \leq M(A).$$

For each $h_\infty > \inf_F \|T_F\|_{\mathcal{H}_\infty}$, a stabilizing state feedback gain achieving $\|T_F\|_{\mathcal{H}_\infty} < h_\infty$ is given by [8, 29]

$$F = \left[I + (1 - h_\infty^{-2})B'XB\right]^{-1} B'XA,$$

where $X \geq 0$ is the stabilizing solution to

$$X = A'X\left[I + (1 - h_\infty^{-2})BB'X\right]^{-1} A, \quad B'XB < h_\infty^2 I.$$

Recall that when system matrix A only has semisimple eigenvalues, the uniform observability of pair $\{\widehat{J}(k)\widetilde{F}, \widetilde{A}\}$ defined in (4.19) is sufficient for the MAS to achieve the state consensus. However, when A contains non-semisimple eigenvalues, the notion of the uniform observability needs to be strengthened. Hence the following assumption is made.

Assumption 10. *The pair $\{\widehat{J}(k)\widetilde{F}, \widetilde{A}\}$ is uniformly observable in the sense that there exists a finite $T_\mathrm{o} > 0$ such that*

$$\sum_{i=0}^{T_\mathrm{o}-1} \left[\widetilde{F}\widetilde{A}^i\right]' \widehat{\mathcal{L}}_m(k+i) \left[\widetilde{F}\widetilde{A}^i\right] \geq \epsilon_\mathrm{o} I \ \ \forall\, k \geq 0 \tag{4.20}$$

and for some $\epsilon_\mathrm{o} > 0$.

The state consensus result is demonstrated below.

Theorem 11. *Under Assumptions* 6, 7, 10*, and the reachability of* (A, B)*, let the state feedback gain be specified by*

$$F = (I + \epsilon_\Delta B'XB)^{-1} B'XA, \tag{4.21}$$

where $\epsilon_\Delta = \epsilon_{\rm o}/\epsilon_{\rm s}$, $\epsilon_{\rm s} = \overline{\sigma}\left\{\sum_{i=0}^{T_{\rm o}-1}\left(\widetilde{F}\widetilde{A}^i\right)'\left(\widetilde{F}\widetilde{A}^i\right)\right\}$*, and* $X \geq 0$ *is the stabilizing solution to*

$$X = A'X\left(I + \epsilon_\Delta BB'X\right)^{-1}A, \quad B'XB < (1-\epsilon_\Delta)^{-1} I. \tag{4.22}$$

Then the feedback homogeneous MAS described in (4.11) *achieves the state consensus defined in* (4.12)*.*

Proof. Since the state consensus is equivalent to the asymptotic stability of the feedback system in Figure 4.2 by Theorem 9, we will show that the given hypotheses ensure such a feedback stability. It is first noted that $0 < 1 - \epsilon_\Delta < 1$. By Lemma 15, the state feedback gain F given by (4.21) implies that the transfer matrix $\widetilde{T}_F(z)$ in Figure 4.2 satisfies $\|\widetilde{T}_F\|_{\mathcal{H}_\infty} = \|T_F\|_{\mathcal{H}_\infty} < 1/\sqrt{1-\epsilon_\Delta}$. The proof can then be completed by showing that $\|\widehat{\Delta}(k)\|_{\mathcal{H}_\infty} \leq \sqrt{1-\epsilon_\Delta}$ in light of Lemma 14. According to Figure 4.1, $\widetilde{y}(k) = \widetilde{F}\widetilde{A}^{(k-k_0)}\widetilde{x}_0$ with $\widetilde{x}_0 \neq 0$ denoting the initial condition at time index k_0. Then it follows from the uniform observability in (4.20) that

$$\begin{aligned}
\sum_{k=k_0}^{k_0+T_{\rm o}-1} \widetilde{y}(k)'\widehat{\mathcal{L}}_m(k)\widetilde{y}(k) &= \widetilde{x}_0'\left\{\sum_{k=k_0}^{k_0+T_{\rm o}-1}\left[\widetilde{F}\widetilde{A}^{(k-k_0)}\right]'\widehat{\mathcal{L}}_m(k)\left[\widetilde{F}\widetilde{A}^{(k-k_0)}\right]\right\}\widetilde{x}_0 \\
&\geq \epsilon_{\rm o}\widetilde{x}_0'\widetilde{x}_0 = \epsilon_\Delta\epsilon_{\rm s}\widetilde{x}_0'\widetilde{x}_0 \\
&\geq \epsilon_\Delta\widetilde{x}_0'\left\{\sum_{k=k_0}^{k_0+T_{\rm o}-1}\left[\widetilde{F}\widetilde{A}^{(k-k_0)}\right]'\left[\widetilde{F}\widetilde{A}^{(k-k_0)}\right]\right\}\widetilde{x}_0 \\
&= \epsilon_\Delta\sum_{k=k_0}^{k_0+T_{\rm o}-1}\|\widetilde{y}(k)\|^2 \quad \forall\ k_0 \geq 0.
\end{aligned}$$

Thus we have

$$\begin{aligned}\sum_{k=k_0}^{k_0+T_o-1} \|\widetilde{v}(k)\|^2 &= \sum_{k=k_0}^{k_0+T_o-1} \widetilde{y}(k)'\widehat{\Delta}(k)^2\widetilde{y}(k) \\ &\leq \sum_{k=k_0}^{k_0+T_o-1} \widetilde{y}(k)'\left[I - \widehat{\mathcal{L}}_m(k)\right]\widetilde{y}(k) \\ &= \sum_{k=k_0}^{k_0+T_o-1} \left[\|\widetilde{y}(k)\|^2 - \widetilde{y}(k)'\widehat{\mathcal{L}}_m(k)\widetilde{y}(k)\right] \\ &\leq (1-\epsilon_\Delta) \sum_{k=k_0}^{k_0+T_o-1} \|\widetilde{y}(k)\|^2 \quad \forall\, k_0 \geq 0,\end{aligned}$$

which implies that $\|\widehat{\Delta}(k)\|_{\mathcal{H}_\infty} \leq \sqrt{1-\epsilon_\Delta}$, thereby concluding the proof. □

Remark 6. The existence of the stabilizing solution to the ARE in (4.22) requires the the stabilizability of (A, B) and the detectability of $(0, A)$ on the unit circle, the latter of which does not hold, given the assumption that all eigenvalues of A lie on the unit circle. An easy fix is to consider computing the stabilizing solution to ARE (4.22) with $Q = \phi I$ added to its right hand side for a sufficiently small $\phi > 0$. This way ensures the detectability of $(Q^{1/2}, A)$ on the unit circle, and the existence of an approximate stabilizing solution to the original ARE. Moreover, since the stabilizing solution X is monotonically increasing with respect to Q, the inequality in (4.22) can always be satisfied as long as $\phi > 0$ is sufficiently small. □

4.2.4 Simulations

To illustrate our design for state consensus control, we use the example of [83, 99] for the MAS consisting of $N = 4$ agents with discretization and modification. Bilinear transform is adopted to convert the continuous-time agent dynamics to the discrete-time agent dynamics, given by

$$A = \begin{bmatrix} 1.0474 & 0.0474 & -0.0524 \\ -0.0524 & 0.9476 & -0.0474 \\ 0.0998 & 0.0998 & 0.9950 \end{bmatrix}, \quad B = \begin{bmatrix} 0.3064 \\ 0.2772 \\ 0.3292 \end{bmatrix},$$

with time-step $T_s = 0.1$. The time-varying communication topology $\mathcal{G}(k)$ of the MAS is shown in Figure 4.3, with the corresponding Laplacian matrix specified by

$$\mathcal{L}(k) = \begin{cases} \mathcal{L}_0, & k = 4\kappa, \\ \mathcal{L}_1, & k = 4\kappa + 1, \\ \mathcal{L}_2, & k = 4\kappa + 2, \\ \mathcal{L}_3, & k = 4\kappa + 3, \end{cases}$$

for $\kappa \in \{0, 1, \ldots\}$, where

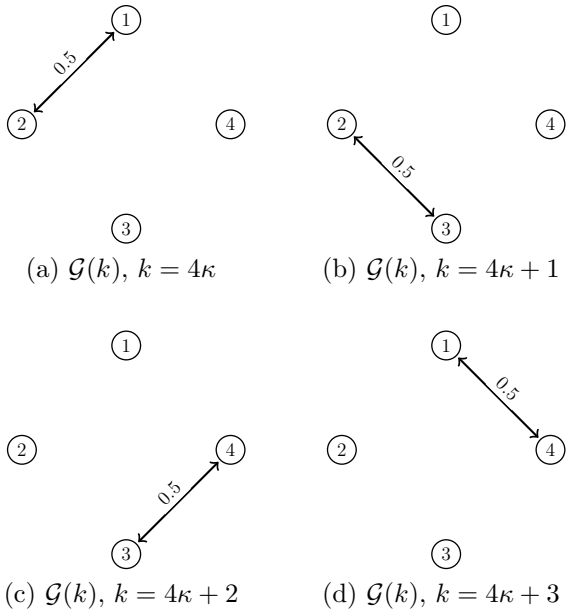
$$\mathcal{L}_0 = 0.5 \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{L}_1 = 0.5 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathcal{L}_2 = 0.5 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad \mathcal{L}_3 = 0.5 \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

Assumption 6 holds since A has three different unstable eigenvalues on the unit circle, i.e., 1, $0.9950 \pm 0.0998j$. Then Assumption 8 can be satisfied by selecting

$$X = \begin{bmatrix} 1.5 & 0.5 & -0.5 \\ 0.5 & 1.5 & 0.5 \\ -0.5 & 0.5 & 1.5 \end{bmatrix}, \quad F = \begin{bmatrix} 0.4635 & 0.7635 & 0.4193 \end{bmatrix}.$$

Note that $\mathcal{G}(k)$ is not connected at any time instant k. However the average of $\mathcal{L}_0$, $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ corresponds to a connected graph, and $\mathcal{L}(k)$ repeats periodically, which indicate that the uniform connectedness imposed in Assumption 7 is satisfied. In addition it is easy to



(a) $\mathcal{G}(k)$, $k = 4\kappa$ (b) $\mathcal{G}(k)$, $k = 4\kappa + 1$

(c) $\mathcal{G}(k)$, $k = 4\kappa + 2$ (d) $\mathcal{G}(k)$, $k = 4\kappa + 3$

Figure 4.3: Time-varying topology with $N = 4$ agent nodes

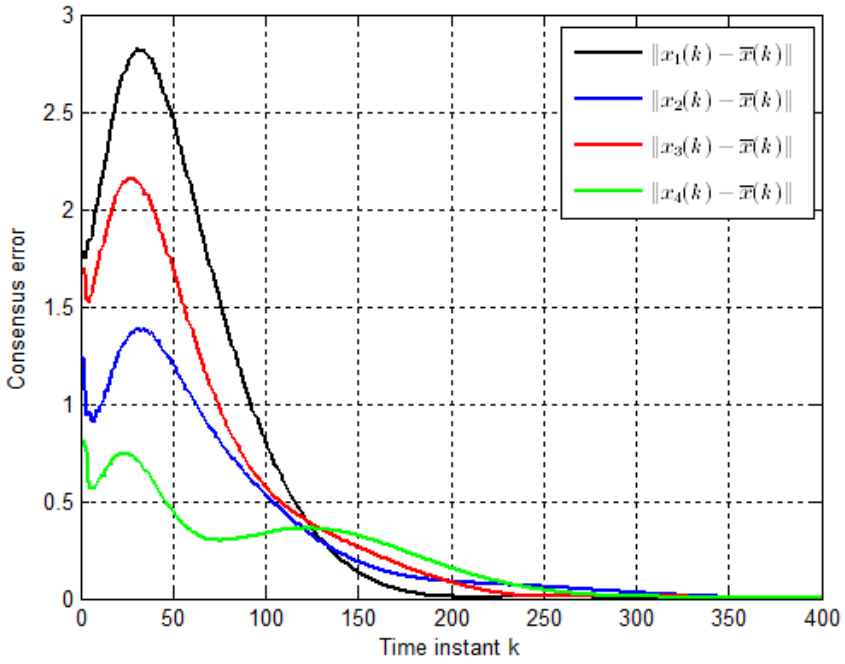


Figure 4.4: Consensus error for each agent

verify that the uniform observability defined in Assumption 9 holds for the given graph $\mathcal{G}(k)$. Therefore the results in Theorem 10 can be applied for state consensus control.

In the simulation, the initial condition $x(0) = x_0$ of the MAS is set as a random vector following Gaussian distribution. The consensus error $\|x_i(k) - \bar{x}(k)\|$ for all the four agents are shown in Figure 4.4, with $\bar{x}(k)$ denoting the average of the four agent state vectors. It can be observed from the plots that the consensus error converges to 0 asymptotically for each $i \in \mathcal{N}$, indicating that state consensus is indeed achieved for the MAS.

4.3 Distributed State Estimation

Chapter 3 is focused on distributed Kalman filtering in the presence of data packet drops, which are modeled as Bernoulli random processes. In this section we study distributed state estimation over WSNs with deterministic time-varying communication topology.

4.3.1 Problem Formulation

A network of N sensor nodes is used to estimate and track the state of a dynamic target system with state space description

$$x(k+1) = Ax(k), \quad x(0) = x_0, \tag{4.23a}$$

$$y(k) = Cx(k), \tag{4.23b}$$

where $x(k) \in \mathbb{R}^n$ is the state vector, and $y(k) \in \mathbb{R}^m$ is the output signal. To lower the communication overhead between the target system and individual sensor nodes, the output signal is required to be transmitted to only one or a few nodes at each time instant. Each node is supposed to locally estimate and track the dynamic state only by using the output signal if received, exchanging information with its neighbors, and carrying out local computation. We thus consider the distributed state estimation protocol

$$\hat{x}_i(k+1) = A\hat{x}_i(k) + Ka_{i,0}(k)\left[y(k) - C\hat{x}_i(k)\right] - K\sum_{j=1}^{N} a_{i,j}(k)C\left[\hat{x}_i(k) - \hat{x}_j(k)\right]$$

for $i \in \mathcal{N}$, where $\hat{x}_i(k)$ is the local estimate of $x(k)$ at the ith sensor node, $a_{i,j}(k) \geq 0$ $\forall\, i, j \in \mathcal{N}$ are the entries of the adjacency matrix associated with time-varying graph $\mathcal{G}(k)$, which represents the interconnection of the sensor network, and $a_{i,0}(k)$ corresponds to the transmission of the output signal. If $y(k)$ is received by node i at time k, then $a_{i,0}(k) > 0$; otherwise $a_{i,0}(k) = 0$. By a slight abuse of notation, we denote $\hat{x}_0(k) = x(k)$. Then the above estimation protocol can be rewritten as

$$\hat{x}_i(k+1) = A\hat{x}_i(k) - KC \sum_{j=0}^{N} a_{i,j}(k) \left[\hat{x}_i(k) - \hat{x}_j(k)\right], \quad i \in \mathcal{N}. \tag{4.24}$$

We now construct an augmented graph $\mathcal{G}_A(k)$ by adding to $\mathcal{G}(k)$ a node v_0, which represents the target system, and edges from v_0 to v_i for all $i \in \mathcal{N}$ satisfying $a_{i,0}(k) > 0$, which mark the information flow from the target system to the ith sensor node. Let $\mathcal{L}(k)$ and $\mathcal{L}_A(k)$ be the Laplacian matrices associated with graph $\mathcal{G}(k)$ and $\mathcal{G}_A(k)$ respectively. It follows that

$$\mathcal{L}_A(k) = \begin{bmatrix} 0 & 0 \\ -\sum\limits_{i=1}^{N} a_{i,0}(k) g_i & \mathcal{M}(k) \end{bmatrix}, \quad \mathcal{M}(k) = \mathcal{L}(k) + \sum_{i=1}^{N} a_{i,0}(k) g_i g_i',$$

where $g_i \in \mathbb{R}^N$ is a vector with 1 in its ith entry and 0 elsewhere. Let $e_{x;i}(k) = x(k) - \hat{x}_i(k)$ be the state estimation error at node i. Taking the difference between (4.23a) and (4.24) yields

$$e_{x;i}(k+1) = Ae_{x;i}(k) - KC \sum_{j=0}^{N} a_{i,j}(k) \left[e_{x;i}(k) - e_{x;j}(k)\right].$$

Denote $e_x(k) = \mathrm{vec}\{e_{x;1}(k), \ldots, e_{x;N}(k)\}$, $\widetilde{A} = I_N \otimes A$, $\widetilde{K} = I_N \otimes K$, $\widetilde{C} = I_N \otimes C$, and $\widetilde{\mathcal{M}}(k) = \mathcal{M}(k) \otimes I_m$. Then the collective estimation error dynamics are given by

$$e_x(k+1) = \left[\widetilde{A} - \widetilde{K}\widetilde{\mathcal{M}}(k)\widetilde{C}\right] e_x(k). \tag{4.25}$$

Similarly denote transfer matrices $G(z) = C(zI - A)^{-1}K$, and $\widetilde{G}(z) = I_N \otimes G(z)$. Then error dynamics (4.25) have the feedback form as configured in Figure 4.5. Hence our goal is to design estimation gain K such that asymptotic stability holds for the time-varying feedback system in Figure 4.5.

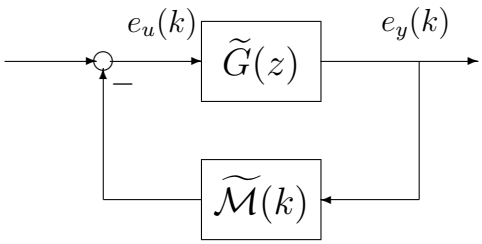


Figure 4.5: Error dynamics (4.25) in feedback form

Similar to section 4.2, the next assumption is made throughout this whole section.

Assumption 11. *All eigenvalues of A are restricted to lie on the unit circle.*

The above assumption is made without loss of generality. On one hand, states governed by stable eigenvalues usually do not require estimation actions since they will exponentially converge to zero as time proceeds; on the other hand, states dominated by unstable eigenvalues outside of the unit circle go exponentially fast to infinity, which are hard and pointless to track in many practical applications. Again we will study distributed state estimation under two different situations: the target system described in (4.23) is neutrally stable, and the target system is neutrally unstable. The corresponding results will be given respectively in the next two subsections.

4.3.2 Neutrally Stable Target System

Recall the definition of uniform connectedness for time-varying graphs. The following restrictions are imposed on communication graph $\mathcal{G}(k)$.

Assumption 12. *The time-varying graph $\mathcal{G}(k)$ with Laplacian matrix $\mathcal{L}(k)$ is undirected, satisfying* (a) $0 \leq \mathcal{M}(k) \leq I \ \forall \ k \geq 0$, *and* (b) $\mathcal{G}(k)$ *is uniformly connected, i.e., there exists a*

finite $T > 0$ such that

$$\overline{\widetilde{\mathcal{M}}}(k) := \frac{1}{T}\sum_{i=0}^{T-1}\widetilde{\mathcal{M}}(k+i) > 0 \ \ \forall\, k \geq 0.$$

For (a), it is noted that $\mathcal{M}(k) = \mathcal{L}(k) + \sum_{i=1}^{N} a_{i,0}(k) g_i g_i' \geq \mathcal{L}(k) \geq 0$ follows from the property of undirected graph, and $\mathcal{M}(k) \leq I$ can be satisfied by appropriate (positive and symmetric) scaling. For (b), considering how we create $\mathcal{G}_A(k)$ based on $\mathcal{G}(k)$, the uniform connectedness of $\mathcal{G}(k)$ implies the uniform connectedness of $\mathcal{G}_A(k)$, i.e.,

$$\overline{\mathcal{L}}_A(k) := \frac{1}{T}\sum_{i=0}^{T-1}\mathcal{L}_A(k+i) = \begin{bmatrix} 0 & 0 \\ -\frac{1}{T}\sum_{i=0}^{T-1}\sum_{j=1}^{N} a_{j,0}(k+i) g_j & \frac{1}{T}\sum_{i=0}^{T-1}\mathcal{M}(k+i) \end{bmatrix}$$

satisfies $|\lambda_2\{\overline{\mathcal{L}}_A(k)\}| > 0 \ \forall\ k \geq 0$. It follows that $\frac{1}{T}\sum_{i=0}^{T-1}\mathcal{M}(k+i) > 0 \ \forall\ k \geq 0$, which is equivalent to the inequality in (b).

The method to find a stabilizing estimation gain is provided next.

Assumption 13. *The pair (C, A) is observable, and there exists matrix $X > 0$ satisfying*

$$\text{(a) } X = AXA', \quad \text{(b) } I - CXC' \geq 0, \quad \text{(c) } K = AXC'.$$

It is important to observe that the existence of X requires A only having semisimple eigenvalues. The way to find a solution X is similar to that in Assumption 8. It is easy to verify that $A - KC$ is a Schur stability matrix. Denote $\widetilde{\Delta}(k) = I - \widetilde{\mathcal{M}}(k)$. Assumption 12 implies that

$$0 \leq \widetilde{\Delta}(k) \leq I \ \ \forall\, k \geq 0.$$

Hence the error dynamics in (4.25) can be rewritten as

$$e_x(k+1) = \left[\left(\widetilde{A} - \widetilde{K}\widetilde{C}\right) + \widetilde{K}\widetilde{\Delta}(k)\widetilde{C}\right] e_x(k),$$

leading to the equivalent feedback system in Figure 4.6 where

$$\widetilde{T}_K(z) = \widetilde{C}\left(zI - \widetilde{A} + \widetilde{K}\widetilde{C}\right)^{-1}\widetilde{K}. \tag{4.26}$$

It is noted that $e_y(k)$ in Figure 4.6 is the same as that in Figure 4.5, while $e_v(k)$ in Figure 4.6 is related to $e_u(k)$ in Figure 4.5 via $e_v(k) = e_u(k) + \widetilde{C}e_x(k)$.

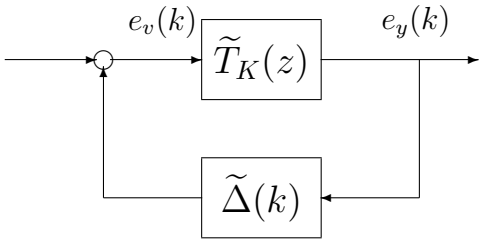


Figure 4.6: Equivalent feedback system to that in Figure 4.5

Lemma 16. *Under Assumption* 11 *and* 13*, the transfer matrix* $G(z)+\frac{1}{2}I = C(zI-A)^{-1}K+\frac{1}{2}I$ *is PR, i.e.,*

$$\left[G(z) + \frac{1}{2}I\right]^* + \left[G(z) + \frac{1}{2}I\right] \geq 0 \ \ \forall\, |z| \geq 1,$$

and (A, K) *is reachable.*

Lemma 17. *Under Assumption* 11 *and* 13*, the transfer matrix* $\widetilde{T}_K(z)$ *in* (4.26) *satisfies* $\|\widetilde{T}_K\|_{\mathcal{H}_\infty} = 1$.

The proofs for Lemma 16 and 17 are similar to those for Lemma 12 and 13 respectively, thus are omitted here. The next result, similar to Lemma 14, is derived directly from the Small Gain Theorem.

Lemma 18. *Consider the closed-loop system configured in Figure* 4.6*. Assume that the transfer matrix* $\widetilde{T}_K(z)$ *is internally stable. Then the feedback system in Figure* 4.6 *is internally stable if* $\|\widetilde{T}_K\|_{\mathcal{H}_\infty}\|\widetilde{\Delta}(k)\|_{\mathcal{H}_\infty} < 1$.

Recall that uniform connectedness is imposed on time-varying graph $\mathcal{G}(k)$ in Assumption 12. However uniformly connected graphs are not adequate to stabilize the error dynamics. Let $\widetilde{J}(k) \geq 0$ such that $\widetilde{\mathcal{M}}(k) = \widetilde{J}(k)^2$. The following assumption is also made.

Assumption 14. *The pair* $\{\widetilde{J}(k)\widetilde{C}, \widetilde{A}\}$ *is uniformly observable in the sense that there exists a finite* $T_{\rm o} > 0$ *such that*

$$\sum_{i=0}^{T_{\rm o}-1} \left[\widetilde{C}\widetilde{A}^i\right]' \widetilde{\mathcal{M}}(k+i) \left[\widetilde{C}\widetilde{A}^i\right] > 0 \ \ \forall\, k \geq 0. \tag{4.27}$$

The solution to distributed state estimation is illustrated below.

Theorem 12. *Under Assumptions* 11 $\sim$ 14, *asymptotic stability holds for the error dynamics in* (4.25) *associated to the distributed state estimation protocol in* (4.24).

Proof. Assumption 11 and 13 implies that $\|\widetilde{T}_K\|_{\mathcal{H}_\infty} = 1$ in light of Lemma 17. The proof can then be completed by showing that $\widetilde{\Delta}(k)$ is a strict contraction map from $\ell_2^m[k_0,\ k_0+T_{\rm o}-1]$ to $\ell_2^m[k_0,\ k_0+T_{\rm o}-1]$ for all $k_0 \geq 0$, in light of Lemma 18. Since $\widetilde{\Delta}(k)$ is symmetric satisfying $0 \leq \widetilde{\Delta}(k) \leq I$ for all $k \geq 0$ by Assumption 12,

$$\begin{aligned}\widetilde{\Delta}(k)^2 &= \left[I - \widetilde{\mathcal{M}}(k)\right]^2 = I - \left[2\widetilde{\mathcal{M}}(k) - \widetilde{\mathcal{M}}(k)^2\right] \leq I - \widetilde{J}(k)\left[2I - \widetilde{\mathcal{M}}(k)\right]\widetilde{J}(k) \\ &\leq I - \widetilde{J}(k)^2 = I - \widetilde{\mathcal{M}}(k) \ \ \forall\, k \geq 0.\end{aligned}$$

It follows that

$$\begin{aligned}\sum_{k=k_0}^{k_0+T_{\rm o}-1} \|e_v(k)\|^2 &= \sum_{k=k_0}^{k_0+T_{\rm o}-1} e_y(k)'\widetilde{\Delta}(k)^2 e_y(k) \\ &\leq \sum_{k=k_0}^{k_0+T_{\rm o}-1} e_y(k)'\left[I - \widetilde{\mathcal{M}}(k)\right] e_y(k) \\ &= \sum_{k=k_0}^{k_0+T_{\rm o}-1} \left[\|e_y(k)\|^2 - e_y(k)'\widetilde{\mathcal{M}}(k) e_y(k)\right] \\ &< \sum_{k=k_0}^{k_0+T_{\rm o}-1} \|e_y(k)\|^2 \quad \forall\, k_0 \geq 0.\end{aligned}$$

The last inequality follows from the uniform observability in (4.27). Indeed this inequality

is violated only if

$$e_y(k)'\widetilde{\mathcal{M}}(k)e_y(k) = 0 \ \ \forall\, k \in [k_0,\ k_0 + T_{\rm o} - 1],$$

where $e_y(k) = \widetilde{C}\widetilde{A}^{(k-k_0)}e_{x0}$ according to Figure 4.5, with $e_{x0} \neq 0$ denoting the initial condition at time index k_0, and hence it contradicts to the uniform observability. The above proves that $\widetilde{\Delta}(k)$ is a strict contraction map from $\ell_2^m[k_0,\ k_0 + T_{\rm o} - 1]$ to $\ell_2^m[k_0,\ k_0 + T_{\rm o} - 1]$ for all $k_0 \geq 0$, thereby concluding the proof. □

4.3.3 Neutrally Unstable Target System

Now we discuss the case where A has at least one non-semisimple eigenvalue. Assumption 12 is imposed again while Assumption 13 fails to hold due to the structure of A. The following lemma, which is dual to Lemma 15, can be useful to the design of estimation gain.

Lemma 19. *Consider $T_K(z) = C(zI - A + KC)^{-1}K$ where $C', K \in \mathbb{R}^{n\times m}$. If the pair (C, A) is detectable, then*

$$M(A)^{\frac{1}{m}} \leq \inf_K \|T_K\|_{\mathcal{H}_\infty} \leq M(A).$$

For each $h_\infty > \inf_K \|T_K\|_{\mathcal{H}_\infty}$, a stabilizing estimation gain achieving $\|T_K\|_{\mathcal{H}_\infty} < h_\infty$ is given by

$$K = AXC'\left[I + (1 - h_\infty^{-2})CXC'\right]^{-1},$$

where $X \geq 0$ is the stabilizing solution to

$$X = A\left[I + (1 - h_\infty^{-2})XC'C\right]^{-1}XA', \quad CXC' < h_\infty^2 I.$$

Recall that when system matrix A only has semisimple eigenvalues, the uniform observability of pair $\{\widetilde{J}(k)\widetilde{C}, \widetilde{A}\}$ defined in (4.27) is sufficient for the estimation error to converge to 0 asymptotically. Nevertheless, when A contains non-semisimple eigenvalues, the notion of the uniform observability needs to be strengthened as follows.

Assumption 15. *The pair* $\{\widetilde{J}(k)\widetilde{C}, \widetilde{A}\}$ *is uniformly observable in the sense that there exists a finite* $T_{\rm o} > 0$ *such that*

$$\sum_{i=0}^{T_{\rm o}-1} \left[\widetilde{C}\widetilde{A}^i\right]' \widetilde{\mathcal{M}}(k+i) \left[\widetilde{C}\widetilde{A}^i\right] \geq \epsilon_{\rm o} I \ \ \forall\, k \geq 0. \tag{4.28}$$

and for some $\epsilon_{\rm o} > 0$.

The solution to distributed state estimation is demonstrated below.

Theorem 13. *Under Assumptions* 11, 12, 15, *and the observability of* (C, A), *let the estimation gain be specified by*

$$K = AXC' \left(I + \epsilon_\Delta CXC'\right)^{-1}, \tag{4.29}$$

where $\epsilon_\Delta = \epsilon_{\rm o}/\epsilon_{\rm s}$, $\epsilon_{\rm s} = \overline{\sigma}\left\{\sum_{i=0}^{T_{\rm o}-1} \left(\widetilde{C}\widetilde{A}^i\right)' \left(\widetilde{C}\widetilde{A}^i\right)\right\}$, *and* $X \geq 0$ *is the stabilizing solution to*

$$X = A\left(I + \epsilon_\Delta XC'C\right)^{-1} XA', \quad CXC' < \left(1 - \epsilon_\Delta\right)^{-1} I. \tag{4.30}$$

Then asymptotic stability holds for the error dynamics in (4.25) *associated to the distributed state estimation protocol in* (4.24).

Proof. It is first noted that $0 < 1 - \epsilon_\Delta < 1$. By Lemma 19, the estimation gain F obtained via (4.29) implies that the transfer matrix $\widetilde{T}_K(z)$ in Figure 4.6 satisfies $\|\widetilde{T}_K\|_{\mathcal{H}_\infty} < 1/\sqrt{1-\epsilon_\Delta}$. The proof can then be completed by showing that $\|\widetilde{\Delta}(k)\|_{\mathcal{H}_\infty} \leq \sqrt{1-\epsilon_\Delta}$ in light of Lemma 18. According to Figure 4.5, $e_y(k) = \widetilde{C}\widetilde{A}^{(k-k_0)} e_{x0}$ with $e_{x0} \neq 0$ denoting the initial condition at time index k_0. Then it follows from the uniform observability in (4.28) that

$$\begin{aligned}
\sum_{k=k_0}^{k_0+T_{\rm o}-1} e_y(k)' \widetilde{\mathcal{M}}(k) e_y(k) &= e'_{x0} \left\{ \sum_{k=k_0}^{k_0+T_{\rm o}-1} \left[\widetilde{C}\widetilde{A}^{(k-k_0)}\right]' \widetilde{\mathcal{M}}(k) \left[\widetilde{C}\widetilde{A}^{(k-k_0)}\right] \right\} e_{x0} \\
&\geq \epsilon_{\rm o} e'_{x0} e_{x0} = \epsilon_\Delta \epsilon_{\rm s} e'_{x0} e_{x0} \\
&\geq \epsilon_\Delta e'_{x0} \left\{ \sum_{k=k_0}^{k_0+T_{\rm o}-1} \left[\widetilde{C}\widetilde{A}^{(k-k_0)}\right]' \left[\widetilde{C}\widetilde{A}^{(k-k_0)}\right] \right\} e_{x0}
\end{aligned}$$

$$= \epsilon_\Delta \sum_{k=k_0}^{k_0+T_o-1} \|e_y(k)\|^2 \quad \forall\ k_0 \geq 0.$$

Thus we have

$$\begin{aligned}
\sum_{k=k_0}^{k_0+T_o-1} \|e_v(k)\|^2 &= \sum_{k=k_0}^{k_0+T_o-1} e_y(k)'\widetilde{\Delta}(k)^2 e_y(k) \\
&\leq \sum_{k=k_0}^{k_0+T_o-1} e_y(k)' \left[I - \widetilde{\mathcal{M}}(k)\right] e_y(k) \\
&= \sum_{k=k_0}^{k_0+T_o-1} \left[\|e_y(k)\|^2 - e_y(k)'\widetilde{\mathcal{M}}(k) e_y(k)\right] \\
&\leq (1-\epsilon_\Delta) \sum_{k=k_0}^{k_0+T_o-1} \|e_y(k)\|^2 \quad \forall\ k_0 \geq 0.
\end{aligned}$$

which implies that $\|\widetilde{\Delta}(k)\|_{\mathcal{H}_\infty} \leq \sqrt{1-\epsilon_\Delta}$, thereby concluding the proof. □

Remark 7. Similar to the situation mentioned in Remark 6, the ARE in (4.30) does not admit a stabilizing solution because $(A, 0)$ is not stabilizable on the unit circle under Assumption 11. An easy way to fix this is to compute the stabilizing solution to ARE (4.30) with $Q = \phi I$ added to its right hand side for a sufficiently small $\phi > 0$. This guarantees the stabilizability of $(A, Q^{1/2})$ on the unit circle, and thus the existence of an approximate stabilizing solution to the original ARE. □

4.3.4 Simulations

To illustrate our design of distributed state estimator, we consider the example of [3] with discretization and modification for a dynamic target system moving in one spatial dimension. The discrete-time system dynamics are given by

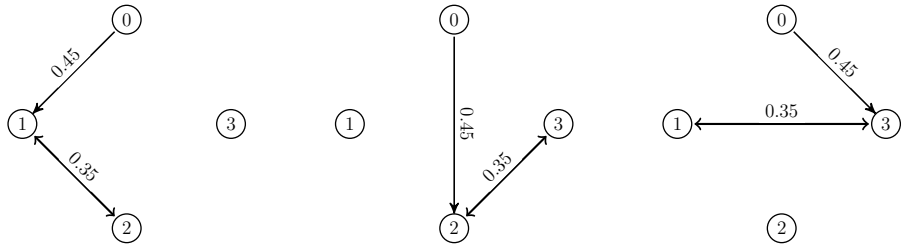
$$A = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

with time-step $T_s = 0.1$. A sensor network of $N = 3$ nodes is employed to estimate and track the state (position and velocity) of the moving target. The time-varying augmented graph $\mathcal{G}_A(k)$ encoding the interconnection between the target node and the sensor nodes is shown in Figure 4.7, with the corresponding $\mathcal{M}(k) = \mathcal{L}(k) + \sum_{i=1}^{N} a_{i,0}(k) g_i g_i'$ specified by

$$\mathcal{M}(k) = \begin{cases} \mathcal{M}_0, & k = 3\kappa, \\ \mathcal{M}_1, & k = 3\kappa + 1, \\ \mathcal{M}_2, & k = 3\kappa + 2, \end{cases}$$

for $\kappa \in \{0, 1, \ldots\}$, where

$$\mathcal{M}_0 = \begin{bmatrix} 0.8 & -0.35 & 0 \\ -0.35 & 0.35 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \mathcal{M}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.8 & -0.35 \\ 0 & -0.35 & -0.35 \end{bmatrix}, \ \mathcal{M}_2 = \begin{bmatrix} 0.35 & 0 & -0.35 \\ 0 & 0 & 0 \\ -0.35 & 0 & 0.8 \end{bmatrix}.$$



(a) $\mathcal{G}_A(k)$, $k = 3\kappa$ (b) $\mathcal{G}_A(k)$, $k = 3\kappa + 1$ (c) $\mathcal{G}_A(k)$, $k = 3\kappa + 2$

Figure 4.7: Time-varying topology with a target node and $N = 3$ sensor nodes

It is important to observe that A has a non-semisimple eigenvalue on the unit circle, i.e., 1. Hence Assumption 11 holds while Assumption 13 cannot be satisfied. Although $\mathcal{G}_A(k)$ is not connected at any time instant k, the average of $\mathcal{M}_0$, $\mathcal{M}_1$ and $\mathcal{M}_2$ is positive definite, and $\mathcal{M}(k)$ repeats periodically, which indicate that the uniform connectedness imposed in Assumption 12 is satisfied. It can also be verified that the uniform observability defined in Assumption 15 holds for $\epsilon_{\mathrm{o}} = 0.3902$ when $T_{\mathrm{o}} = 18$. Therefore the results in Theorem 13 can be applied to design the distributed state estimator. By direct calculations, $\epsilon_{\mathrm{s}} = 33.2252$,

and thus $\epsilon_\Delta = 0.0117$. Then the approximate stabilizing solution to the ARE in (4.30) and the resulting estimation gain are given respectively by

$$X = \begin{bmatrix} 1.0073 & 0.0588 \\ 0.0588 & 0.0069 \end{bmatrix}, \quad K = \begin{bmatrix} 1.0014 \\ 0.0581 \end{bmatrix}.$$

In the simulation, the initial condition $e_x(0) = e_{x0}$ is set as a random vector following Gaussian distribution. The estimation error $\|e_{x;i}(k)\|$ at all the three sensor nodes are presented in Figure 4.8. It is revealed from the plots that the estimation error converges to 0 asymptotically for each $i \in \mathcal{N}$, indicating a close tracking of the dynamic state.

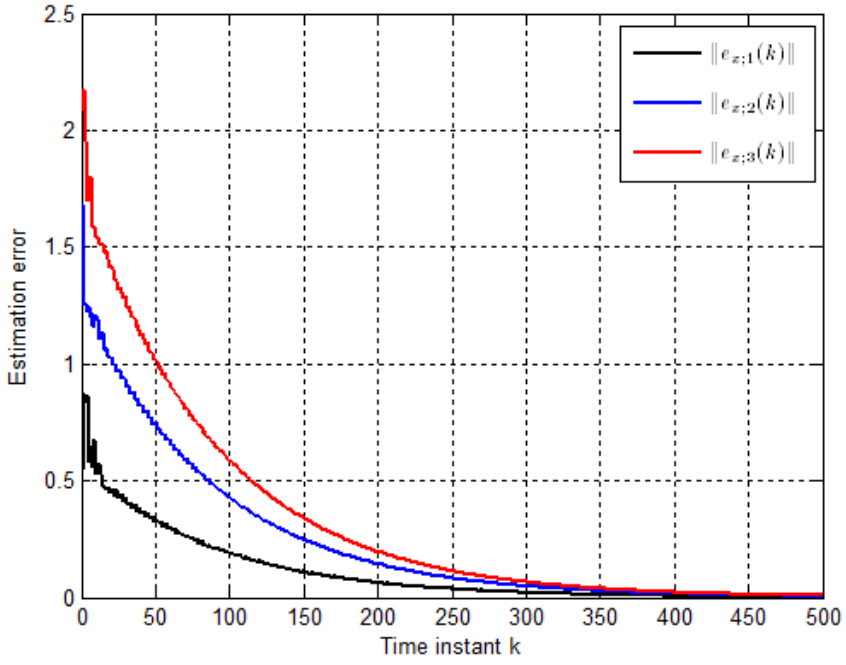


Figure 4.8: Estimation error at each sensor node

CHAPTER 5
CONCLUSION AND FUTURE WORK

5.1 Conclusion

Several basic but important issues concerning Kalman filtering over the WSN are studied in this dissertation. Chapter 2 is focused on data fusion Kalman filtering for discrete-time LTI systems in the presence of data packet drops. First we consider the case where only one sensor node is employed to transmit its measurement in a single packet to the data fusion center. Building on the mathematical foundations established in [77] and treating packet drop as a special case of multiplicative noises, we show that the widely studied critical arrival rate of the Bernoulli packet drop channel can be computed by solving a simple LMI problem. Then we consider a more general scenario where multiple sensor nodes are used to obtain and transmit observations to the fusion center through different packet drop channels. Under the TCP-like protocol, the stationary Kalman filter that minimizes the average error variance is studied in the form of an one-step predictor, and the optimal estimation gain is derived in terms of the stabilizing solution to the corresponding MARE. The MARE admits the stabilizing solution when the stability margin, which can be computed by solving a set of LMIs, is greater than or equal to one.

Chapter 3 is focused on distributed Kalman filtering for discrete-time LTI systems, where each node in the WSN is required to locally estimate the dynamic state in a collaborative manner with its neighbors in the presence of data packet drops. We first present the stationary DKF that minimizes the local average error variance in the steady-state at each sensor node, based on the stabilizing solution to the corresponding MARE. The stability issue is again addressed by adopting the stability margin. Following [55], we then propose the KCF by combining the stationary DKF with a consensus term of prior estimates to achieve consensus among different local estimates. The upper bound for the consensus coefficient can be computed by solving a simple LMI problem to ensure the MS stability of the estimation error

dynamics associated to the KCF. Simulations show that the proposed KCF outperforms the stationary DKF in general.

In Chapter 4 we consider the situation where the communication between a network of interconnected nodes is represented by a deterministic time-varying topology. First we study state consensus control for discrete-time homogeneous MASs over deterministic time-varying feedback topology. The proposed solutions generalize the restrictive results for state consensus in [92] and [38], and can be applied to MASs of more general form. Then we investigate distributed state estimation over the WSN with deterministic time-varying communication topology. This problem is addressed for neutrally stable target systems and neutrally unstable target systems respectively. The proposed estimation protocol allows low communication overhead between the target system and distributed sensor nodes since the observation information is required to be transmitted to only one or a few nodes at each time instant. Each sensor node can keep a close track of the target state in a collaborative manner with its neighbors as long as uniform observability holds for the time-varying graph.

5.2 Future Work

In this section we highlight our points of view for future research work.

There are several technical issues in this dissertation that require our further consideration.

- In Chapter 2 we show that the critical arrival rate can be computed by solving a simple LMI problem. It is easy to see that when measurement $y(k)$ has dimension one, i.e., $m = 1$, the critical arrival rate $p_{\inf}$ is given by $p_{\inf} = 1 - M(A)^{-2}$. However the closed-form expression for $p_{\inf}$ remains unknown when $m > 1$.

- Data packet drops are considered as a special case of fading channels and treated as multiplicative noises. Then the necessary and sufficient condition for the MS stability of feedback systems over fading channels is adopted to tackle the MS stabilizability of the estimation error dynamics associated to the data fusion Kalman filter, and the existence of

the stabilizing solution to MARE. It is reasonable and useful to consider how to generalize these results for Kalman filtering over general fading channels instead of packet drop channels. This problem can be quite challenging, especially when different fading channels are correlated to each other.

- In Chapter 3 the $\mathcal{H}_\infty$ norm of stochastic systems is adopted in pursuit of an upper bound for the consensus coefficient ε of KCF. A drawback of the proposed solution is its conservativeness. As pointed out in the simulation examples, MS stability may still holds for the error dynamics associated to the KCF even if the selected ε is greater than the upper bound obtained via Theorem 6 or Corollary 1. Hence more work needs to be done to find a less conservative upper bound for ε.

- Our proposed KCF consists of the stationary DKF and a consensus term of prior estimates. For future work we can consider another widely studied strategy in which the stationary DKF is implemented first, followed by one extended step where each sensor node merges its local estimate with those from its neighbors by a weighted average approach. In this case the weights need to be optimized to yield minimum error variance at each sensor node. The stability will be more difficult to analyze since two separate steps are included in the filtering process. We can also compare this new filter with KCF to see which one has a better performance.

- In Chapter 4 the distributed state estimation problem is addressed for neutrally unstable target systems. It is claimed that the strengthened uniform observability in (4.28) has to be satisfied, and the knowledge of ϵ_{o} value is essential to the design of estimation gain. Although ϵ_{o} can be easily computed in our simulation example, it is hard to be obtained in many practical cases. A possible solution is to adaptively estimate the value of ϵ_{o} using real-time data. More work has to be done to establish a complete answer.

- Robust filtering can be studied for state space model with uncertainty. In addition, we can consider the case where observations obtained at a set of consecutive time instants

are grouped together in a batch for transmission. This problem can be investigated under the framework of multirate system. Another direction for future work is filtering in the presence of time delays. A preliminary solution can be established based on the well-known Smith predictor.

Standard Kalman filtering is only applicable to linear systems. However most of the physical systems are nonlinear in nature. Hence there is a need to study filtering algorithms for nonlinear systems. A large number of nonlinear filters have been proposed over the years [85]. Some are fairly general while other are tailored to particular applications. Widely studied nonlinear filters can be grouped into three broad categories [71]: (i) analytic approximations, (ii) Gaussian sum filters, and (iii) sampling approaches. Denote $p[x(k)|\mathcal{Y}(k)]$ as the posterior probability density function (PDF) of the state $x(k)$ based on the sequence of measurements $\mathcal{Y}(k) := \{y(t)\}_{t=0}^{k}$.

- The extended Kalman filter introduced in Chapter 1 is referred to as analytic approximations because the approximations (linearization) of the nonlinear functions in the state dynamics and measurement model are conducted analytically. Notice that EKF always approximates PDF $p[x(k)|\mathcal{Y}(k)]$ to be Gaussian. If the system model possesses severe nonlinearity, the non-Gaussianity of $p[x(k)|\mathcal{Y}(k)]$ will be more pronounced, and consequently the performance of EKF will be degraded tremendously.

- The main feature of Gaussian sum filters [80] is that they approximate the posterior density $p[x(k)|\mathcal{Y}(k)]$ by a weighted sum of different Gaussian density functions. The approximation can be made as accurate as desirable via the choice of the total number of different functions. The difficulty lies in the real-time computation of weights, means, and covariances. This type of approximation poses great advantage over the other two types, when the posterior density is multimodal.

- The sampling approach is adopted in the unscented Kalman filter (UKF) [87] and the particle filter [6]. The UKF approximates the posterior PDF $p[x(k)|\mathcal{Y}(k)]$ by a Gaussian

density, which is represented by a small number of deterministically chosen sample points. The particle filter embraces a similar scenario, but uses a large number of random (Monte Carlo) sample points. As a result, fairly high computation power is required for the implementation of the particle filter.

It will be interesting to extend our research on the implementation and applications of nonlinear filtering over WSNs by studying these nonlinear filters.

Future work can also be focused on practical applications of Kalman filtering over WSN. As introduced in Chapter 1, Kalman filtering over WSN is extensively used for the localization and tracking of moving targets. Another important application is state estimation for smart grids [100, 66, 2], which claims a significant position in modern power system operations. State estimation for smart grids usually serves as a critical prerequisite for other operation functions such as real-time monitoring, load forecasting, frequency control, etc. Static state estimators are widely adopted in smart grids due to the modest traditional monitoring technology. Nowadays with the rapid development of phasor measurement technology, there has been an explosion in the use of dynamic estimators for real-time monitoring and control of highly complex and dynamic power systems.

Kalman filtering is one of the most popular dynamic state estimation techniques that can recursively compute the optimal state estimate for smart grids, given the true state and measurement models. However the exact noise statistics can hardly be obtained, and measurement errors often occur due to device failure or malicious data attacks. In [100] an adaptive Kalman filter with inflatable noise variances is proposed for smart grid state estimation with real-time voltage phasor measurements, which demonstrates remarkable advantage in dealing with various adverse conditions including inaccurate system models and measurement errors.

In addition, most of the power systems in the real world are nonlinear. Hence the implementation of nonlinear filters such as EKF and UKF in smart grids is also widely studied in existing literatures.

Recently, due to the concerns regarding global warming and energy crisis, people start to integrate renewable distributed energy resources (DERs) such as solar and wind power into the smart grid in their research. Since the power generation patterns of DERs are mostly intermittent in nature and distributed over the grid, WSN provides a feasible and cost-effective sensing and communication solution for smart grid operations. In [66] a set of sensor nodes are deployed to get observations on DER, which are then transmitted to the nearby base station. An accuracy-dependent Kalman filter is proposed to estimate the DER states, which are then fed back to the DER for control purpose.

As a conclusion, there are many potential research directions in this area worthy of our exploration in the near future.

REFERENCES

[1] P. Alriksson and A. Rantzer. Distributed Kalman filtering using weighted averaging. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 2445–2450, 2006.

[2] L. Alvergue, B. Tang, and G. Gu. A consensus approach to dynamic state estimation for smart grid power systems. In *Intelligent Control and Automation (WCICA), 2014 11th World Congress on*, pages 174–179. IEEE, 2014.

[3] L. D. Alvergue, A. Pandey, G. Gu, and X. Chen. Consensus control for heterogeneous multiagent systems. *SIAM Journal on Control and Optimization*, 54(3):1719–1738, 2016.

[4] B. D. O. Anderson and J. B. Moore. Optimal filtering. *Englewood Cliffs*, 21:22–95, 1979.

[5] B. D. O. Anderson and S. Vongpanitlerd. *Network analysis and synthesis*. Dover, 2006.

[6] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002.

[7] B. Bamieh. Structured stochastic uncertainty. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 1498–1503. IEEE, 2012.

[8] T. Basar and P. Bernhard. $\mathcal{H}_\infty$ optimal control and related minimax design problems. *Birkhaüser, Boston, Massachusetts*, 1991.

[9] R. Bellman. *Introduction to matrix analysis*. SIAM, 1997.

[10] B. Bollobás. *Modern graph theory*, volume 184. Springer Science & Business Media, 2013.

[11] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*, volume 290. Citeseer, 1976.

[12] J. A. Bondy and U. S. R. Murty. Graph theory (graduate texts in mathematics), 2008.

[13] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[14] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg. Feedback stabilization over signal-to-noise ratio constrained channels. *Automatic Control, IEEE Transactions on*, 52(8):1391–1403, 2007.

[15] M. Castillo-Effer, D. H. Quintela, W. Moreno, R. Jordan, and W. Westhoff. Wireless sensor networks for flash-flood alerting. In *Devices, Circuits and Systems, 2004. Proceedings of the Fifth IEEE International Caracas Conference on*, volume 1, pages 142–146. IEEE, 2004.

[16] F. S. Cattivelli and A. H. Sayed. Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Transactions on automatic control*, 55(9):2069–2084, 2010.

[17] A. Censi. Kalman filtering with intermittent observations: convergence for semi-Markov chains and an intrinsic performance measure. *Automatic Control, IEEE Transactions on*, 56(2):376–381, 2011.

[18] F. R. K. Chung. Spectral graph theory (CBMS regional conference series in mathematics, no. 92). 1996.

[19] P. Deng and P. Z. Fan. An AOA assisted TOA positioning system. In *Communication Technology Proceedings, 2000. WCC-ICCT 2000. International Conference on*, volume 2, pages 1501–1504. IEEE, 2000.

[20] A. El Bouhtouri, D. Hinrichsen, and A. J. Pritchard. $\mathcal{H}^\infty$-type control for discrete-time stochastic systems. *International Journal of Robust and Nonlinear Control*, 9(13):923–948, 1999.

[21] A. El Bouhtouri, D. Hinrichsen, and A. J. Pritchard. Stability radii of discrete-time stochastic systems with respect to blockdiagonal perturbations. *Automatica*, 36(7):1033–1040, 2000.

[22] N. Elia. Remote stabilization over fading channels. *Systems & Control Letters*, 54(3):237–249, 2005.

[23] E. O. Elliott. Estimates of error rates for codes on burst-noise channels. *Bell system technical journal*, 42(5):1977–1997, 1963.

[24] T. Gao, D. Greenspan, M. Welsh, R. R. Juang, and A. Alm. Vital signs monitoring and patient tracking over a wireless network. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 102–105. IEEE, 2006.

[25] A. Gasparri, S. Panzieri, F. Pascucci, and G. Ulivi. An interlaced extended Kalman filter for sensor networks localisation. *International Journal of Sensor Networks*, 5(3):164–172, 2009.

[26] E. N. Gilbert. Capacity of a burst-noise channel. *Bell system technical journal*, 39(5):1253–1265, 1960.

[27] T. Glad and L. Ljung. *Control theory*. CRC press, 2000.

[28] G. H. Golub and C. F. Van Loan. Matrix computations. 1996. *Johns Hopkins University, Press, Baltimore, MD, USA*, pages 374–426, 1996.

[29] M. Green and D. J. N. Limebeer. *Linear robust control*. Courier Corporation, 2012.

[30] G. Gu, X. Cao, and H. Badr. Generalized LQR control and Kalman filtering with relations to computations of inner-outer and spectral factorizations. *IEEE transactions on automatic control*, 51(4):595–605, 2006.

[31] G. Gu, L. Marinovici, and F. L. Lewis. Consensusability of discrete-time dynamic multiagent systems. *IEEE Transactions on Automatic Control*, 57(8):2085–2089, 2012.

[32] G. Gu and L. Qiu. Stabilization of networked multi-input systems with channel resource allocation. In *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*, pages 59–63. IEEE, 2008.

[33] G. Gu and L. Qiu. Networked control systems for multi-input plants based on polar logarithmic quantization. *Systems & Control Letters*, 69:16–22, 2014.

[34] A. Harter, A. Hopper, P. Steggles, A. Ward, and P. Webster. The anatomy of a context-aware application. *Wireless Networks*, 8(2/3):187–197, 2002.

[35] K. C. Ho and Y. T. Chan. Solution and performance analysis of geolocation by TDOA. *IEEE Transactions on Aerospace and Electronic Systems*, 29(4):1311–1322, 1993.

[36] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 2012.

[37] M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica*, 43(4):598–607, 2007.

[38] A. Jadbabaie, J. Lin, and S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on automatic control*, 48(6):988–1001, 2003.

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

[40] R. Khan, S. U. Khan, S. Khan, and M. U. A. Khan. Localization performance evaluation of extended Kalman filter in wireless sensors network. *Procedia Computer Science*, 32:117–124, 2014.

[41] U. A. Khan and J. M. F. Moura. Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, 56(10):4919–4935, 2008.

[42] G. Köthe. *Topological vector spaces*. Springer, 1983.

[43] W. Li, Y. Jia, and J. Du. Distributed Kalman consensus filter with intermittent observations. *Journal of the Franklin Institute*, 352(9):3764–3781, 2015.

[44] N. P. Mahalik. *Sensor Networks and Configuration*. Springer, 2007.

[45] R. Merris. A survey of graph Laplacians. *Linear and Multilinear Algebra*, 39(1-2):19–31, 1995.

[46] Y. Mo and B. Sinopoli. A characterization of the critical value for Kalman filtering with intermittent observations. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 2692–2697. IEEE, 2008.

[47] Y. Mo and B. Sinopoli. Kalman filtering with intermittent observations: Tail distribution and critical value. *IEEE Transactions on Automatic Control*, 57(3):677–689, 2012.

[48] P. Mohan, V. N. Padmanabhan, and R. Ramjee. Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 323–336. ACM, 2008.

[49] L. Moreau. Stability of continuous-time distributed consensus algorithms. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 4, pages 3998–4003. IEEE, 2004.

[50] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on automatic control*, 50(2):169–182, 2005.

[51] D. Niculescu and B. Nath. Ad hoc positioning system (APS) using AOA. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 3, pages 1734–1743. Ieee, 2003.

[52] P. Ogren, E. Fiorelli, and N. E. Leonard. Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment. *IEEE Transactions on Automatic control*, 49(8):1292–1302, 2004.

[53] R. Olfati-Saber. Distributed Kalman filter with embedded consensus filters. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 8179–8184. IEEE, 2005.

[54] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Decision and Control, 2007 46th IEEE Conference on*, pages 5492–5498. IEEE, 2007.

[55] R. Olfati-Saber. Kalman-consensus filter: Optimality, stability, and performance. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 7036–7042. IEEE, 2009.

[56] R. Olfati-Saber and R. M. Murray. Graph rigidity and distributed formation stabilization of multi-vehicle systems. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 3, pages 2965–2971. IEEE, 2002.

[57] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on automatic control*, 49(9):1520–1533, 2004.

[58] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov. System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of mobile multimedia*, 1(4):307–326, 2006.

[59] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal. Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal processing magazine*, 22(4):54–69, 2005.

[60] N. Patwari and A. O. Hero. Using proximity and quantized RSS for sensor localization in wireless networks. In *Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, pages 20–29. ACM, 2003.

[61] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'dea. Relative location estimation in wireless sensor networks. *IEEE Transactions on signal processing*, 51(8):2137–2148, 2003.

[62] K. Plarre and F. Bullo. On Kalman filtering for detectable systems with intermittent observations. *Automatic Control, IEEE Transactions on*, 54(2):386–390, 2009.

[63] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43. ACM, 2000.

[64] L. Qiu, G. Gu, and W. Chen. Stabilization of networked multi-input systems with channel resource allocation. *IEEE Transactions on Automatic Control*, 58(3):554–568, 2013.

[65] H. J. Rad, T. Van Waterschoot, and G. Leus. Cooperative localization using efficient Kalman filtering for mobile wireless sensor networks. In *Signal Processing Conference, 2011 19th European*, pages 1984–1988. IEEE, 2011.

[66] M. M. Rana and L. Li. An overview of distributed microgrid state estimation and control for smart grids. *Sensors*, 15(2):4302–4325, 2015.

[67] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on automatic control*, 50(5):655–661, 2005.

[68] W. Ren and R. W. Beard. *Distributed consensus in multi-vehicle cooperative control*. Springer, 2008.

[69] W. Ren, R. W. Beard, and E. M. Atkins. A survey of consensus problems in multiagent coordination. In *American Control Conference, 2005. Proceedings of the 2005*, pages 1859–1864. IEEE, 2005.

[70] W. Ren, R. W. Beard, and D. B. Kingston. Multi-agent Kalman consensus with relative uncertainty. In *American Control Conference, 2005. Proceedings of the 2005*, pages 1865–1870. IEEE, 2005.

[71] B. Ristic, S. Arulampalam, and N. J. Gordon. *Beyond the Kalman filter: Particle filters for tracking applications*. Artech house, 2004.

[72] E. R. Rohr, D. Marelli, and M. Fu. Kalman filtering with intermittent observations: on the boundedness of the expected error covariance. 2013.

[73] A. Savvides, C. Han, and M. B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 166–179. ACM, 2001.

[74] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163–187, 2007.

[75] P. Seiler and R. Sengupta. Analysis of communication losses in vehicle control problems. In *American Control Conference, 2001. Proceedings of the 2001*, volume 2, pages 1491–1496. IEEE, 2001.

[76] A. Shareef and Y. Zhu. Localization using extended Kalman filters in wireless sensor networks. *Kalman Filter: Recent Advances and Applications*, pages 297–320, 2009.

[77] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. Kalman filtering with intermittent observations. *IEEE transactions on Automatic Control*, 49(9):1453–1464, 2004.

[78] A. Smith, H. Balakrishnan, M. Goraczko, and N. Priyantha. Tracking moving devices with the cricket location system. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 190–202. ACM, 2004.

[79] E. Song, J. Xu, and Y. Zhu. Optimal distributed Kalman filtering fusion with singular covariances of filtering errors and measurement noises. *IEEE Transactions on Automatic Control*, 59(5):1271–1282, 2014.

[80] H. W. Sorenson and D. L. Alspach. Recursive Bayesian estimation using Gaussian sums. *Automatica*, 7(4):465–479, 1971.

[81] J. M. Steele. *The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities*. Cambridge University Press, 2004.

[82] G. L. Stüber. *Principles of mobile communication*. Springer Science & Business Media, 2011.

[83] Y. Su and J. Huang. Stability of a class of linear switching systems with applications to two consensus problems. *IEEE Transactions on Automatic Control*, 57(6):1420–1430, 2012.

[84] H. Suzuki. A statistical model for urban radio propogation. *IEEE Transactions on communications*, 25(7):673–680, 1977.

[85] H. Tanizaki. *Nonlinear filters: estimation and applications*. Springer Science & Business Media, 2013.

[86] P. Vaidyanathan. The discrete-time bounded-real lemma in digital filtering. *IEEE transactions on circuits and systems*, 32(9):918–924, 1985.

[87] E. A. Wan and R. Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. IEEE, 2000.

[88] S. Wang, W. Ren, and Z. Li. Information-driven fully distributed Kalman filter for sensor networks in presence of naive nodes. *arXiv preprint arXiv:1410.0411*, 2014.

[89] X. Wang, M. Fu, and H. Zhang. Target tracking in wireless sensor networks based on the combination of KF and MLE using distance measurements. *IEEE Transactions on Mobile Computing*, 11(4):567–576, 2012.

[90] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Transactions on Information Systems (TOIS)*, 10(1):91–102, 1992.

[91] C. Xiao and D. J. Hill. Generalizations and new proof of the discrete-time positive real lemma and bounded real lemma. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 46(6):740–743, 1999.

[92] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th international symposium on Information processing in sensor networks*, page 9. IEEE Press, 2005.

[93] N. Xiao, L. Xie, and L. Qiu. Feedback stabilization of discrete-time networked systems over fading channels. *Automatic Control, IEEE Transactions on*, 57(9):2176–2189, 2012.

[94] L. Xie and L. Xie. Stability of a random Riccati equation with Markovian binary switching. *Automatic Control, IEEE Transactions on*, 53(7):1759–1764, 2008.

[95] K. Xu, J. Zhou, F. Liu, and G. Gu. LQG control for networked control systems in the presence of data packet drops. In *Control Conference (CCC), 2016 35th Chinese*, pages 7428–7433. IEEE, 2016.

[96] K. You, M. Fu, and L. Xie. Mean square stability for Kalman filtering with Markovian packet losses. *Automatica*, 47(12):2647–2657, 2011.

[97] Z. R. Zaidi and B. L. Mark. Real-time mobility tracking algorithms for cellular networks based on Kalman filtering. *IEEE Transactions on Mobile Computing*, 4(2):195–208, 2005.

[98] F. Zhang. *The Schur complement and its applications*, volume 4. Springer Science & Business Media, 2006.

[99] J. Zhang, G. Gu, and X. Chen. State consensus for homogeneous multi-agent systems over time-varying undirected graph. In *Control and Automation (ICCA), 2016 12th IEEE International Conference on*, pages 31–36. IEEE, 2016.

[100] J. Zhang, G. Welch, N. Ramakrishnan, and S. Rahman. Kalman filters for dynamic and secure smart grid state estimation. *Intelligent Industrial Systems*, 1(1):29–36, 2015.

[101] M. Zhang and T. Tarn. Hybrid control of the pendubot. *IEEE/ASME transactions on mechatronics*, 7(1):79–86, 2002.

[102] J. Zheng, W. Chen, L. Shi, and L. Qiu. Linear quadratic optimal control for discrete-time LTI systems with random input gains. In *Control Conference (CCC), 2012 31st Chinese*, pages 5803–5808. IEEE, 2012.

[103] J. Zheng and L. Qiu. Infinite-horizon linear quadratic optimal control for discrete-time LTI systems with random input gains. In *American Control Conference (ACC), 2013*, pages 1195–1200. IEEE, 2013.

[104] K. Zhou and J. C. Doyle. *Essentials of robust control*, volume 104. Prentice hall Upper Saddle River, NJ, 1998.

VITA

Jianming Zhou was born in Shanghai, China. He is the son of Zemin Zhou and Wenying Zhang. He completed his undergraduate program of Automation and received the degree of Bachelor of Science in Engineering at Shanghai Jiao Tong University, Shanghai, China, in June 2013. In Fall 2013, he came to the United States and joined the Division of Electrical and Computer Engineering at Louisiana State University, Baton Rouge, in pursuit of his graduate study, focusing on Systems Control. He was awarded the degree of Master of Science in Electrical Engineering by Louisiana State University in May 2015. He is currently a candidate for the degree of Doctor of Philosophy in Electrical Engineering to be awarded by Louisiana State University in August 2017.