1-28-2015

Information Theory and Cooperative Control in Networked Multi-Agent Systems with Applications to Smart Grid

Yongxiang Ruan

Yongxiang Ruan

Candidate

Electrical & Computer Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Sudharman K. Jayaweera , Chairperson

Dr. Chaouki T. Abdallah

Dr. Gregory L. Heileman

Dr. Francesco Sorrentino

Information Theory and Cooperative Control in Networked Multi-Agent Systems with Applications to Smart Grid

by

Yongxiang Ruan

Bachelor, Automation, Huazhong University of Sci. & Tech., 2006
M.S., Electrical Engineering, University of New Mexico, 2009

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2014

Dedication

I would like to dedicate this work to my God as a christian and follower of Christ. The very faith in Christ kept me in the program when I intended to transfer, almost dropped out of the program and even out of school. The very faith in Christ sustained me throughout the completion of this Ph.D. program. The very faith in Christ encouraged, comforted and delivered me when I was deep in valleys and in all of my troubles. The very faith in Christ gave me hope and promise in finishing this Ph.D. even if it was dark and hopeless in my sight.

"The LORD is my rock, and my fortress, and my deliverer; my God, my strength, in whom I will trust; my buckler, and the horn of my salvation, and my high tower." –Psalms 18:2.

Acknowledgments

I wholeheartedly appreciate my advisor Dr. Sudharman K. Jayaweera for his long-lasting patience and immense mercy and grace to me which kept me on track when I kept failing and made the completion of this dissertation possible.

I also thank him for his continuous guidance in my intermittent study to help me conduct research work and complete each project in this dissertation.

Besides, I would like to give special thanks to my pastor Rebecca Chan, brother Zhanliang Sun and sister Feng Cheng for their helping hands to me in my desperate situations. They lifted me out of the slimy pit, out of the mud and set my feet on a rock, gave me a firm place to stand (Psalm 40:2).

I also give my thanks to my loving parents Jianming Ruan and Miaoxia Lu. It is the effort of a whole family in nearly every aspect throughout the completion of this program, i.e. finance, daily life care, time-spending and encouragement.

At last, I thank my two aunts who kept encouraging me and caring for me during the entire process of this long-term Ph.D. program, which is a great support to me.

SUBMITTED BY: Yongxiang Ruan

SUPERVISOR: Dr. Sudharman K. Jayaweera
Department of Electrical and Computer Engineering

COMMITTEE MEMBERS: Dr. Chaouki T. Abdallah
Department of Electrical and Computer Engineering

Dr. Gregory L. Heileman
Department of Electrical and Computer Engineering

Dr. Francesco Sorrentino
Department of Mechanical Engineering

Information Theory and Cooperative Control in Networked Multi-Agent Systems with Applications to Smart Grid

by

Yongxiang Ruan

Bachelor, Automation, Huazhong University of Sci. & Tech., 2006
M.S., Electrical Engineering, University of New Mexico, 2009

Ph.D., Engineering, University of New Mexico, 2014

Abstract

This dissertation focuses on information theoretic aspects of and cooperative control techniques in networked multi-agent systems (NMAS) with communication constraints.

In the first part of the dissertation, information theoretic limitations of tracking problems in networked control systems, especially leader-follower systems with communication constraints, are studied. Necessary conditions on the data rate of each communication link for tracking of the leader-follower systems are provided. By considering the forward and feedback channels as one cascade channel, we also provide a lower bound for the data rate of the cascade channel for the system to track a reference signal such that the tracking error has finite second moment. Finally, the aforementioned results are extended to the case in which the leader system and follower system have different system models.

In the second part, we propose an easily scalable hierarchical decision-making and control architecture for smart grid with communication constraints in which distributed customers equipped

with renewable distributed generation (RDG) interact and trade energy in the grid. We introduce the key components and their interactions in the proposed control architecture and discuss the design of distributed controllers which deal with short-term and long-term grid stability, power load balancing and energy routing. At microgrid level, under the assumption of user cooperation and inter-user communications, we propose a distributed networked control strategy to solve the demand-side management problem in microgrids. Moreover, by considering communication delays between users and microgrid central controller, we propose a distributed networked control strategy with prediction to solve the demand-side management problem with communication delays.

In the third part, we consider the disturbance attenuation and stabilization problem in networked control systems. To be specific, we consider the string stability in a large group of interconnected systems over a communication network. Its potential applications could be found in formation tracking control in groups of robots, as well as uncertainty reduction and disturbance attenuation in smart grid. We propose a leader-following consensus protocol for such interconnected systems and derive the sufficient conditions, in terms of communication topology and control parameters, for string stability. Simulation results and performance in terms of disturbance propagation are also given.

In the fourth part, we consider distributed tracking and consensus in networked multi-agent systems with noisy time-varying graphs and incomplete data. In particular, a distributed tracking with consensus algorithm is developed for the space-object tracking with a satellite surveillance network. We also intend to investigate the possible application of such methods in smart grid networks. Later, conditions for achieving distributed consensus are discussed and the rate of convergence is quantified for noisy time-varying graphs with incomplete data. We also provide detailed simulation results and performance comparison of the proposed distributed tracking with consensus algorithm in the case of space-object tracking problem and that of distributed local Kalman filtering with centralized fusion and centralized Kalman filter.

The information theoretic limitations developed in the first part of this dissertation provide

guildlines for design and analysis of tracking problems in networked control systems. The results reveal the mutual interaction and joint application of information theory and control theory in networked control systems. Second, the proposed architectures and approaches enable scalability in smart grid design and allow resource pooling among distributed energy resources (DER) so that the grid stability and optimality is maintained. The proposed distributed networked control strategy with prediction provides an approach for cooperative control at RDG-equipped customers within a self-contained microgrid with different feedback delays. Our string stability analysis in the third part of this dissertation allows a single networked control system to be extended to a large group of interconnected subsystems while system stability is still maintained. It also reveals the disturbance propagation through the network and the effect of disturbance in one subsystem on other subsystems. The proposed leader-following consensus protocol in the constrained communication among users reveals the effect of communication in stabilization of networked control systems and the interaction between communication and control over a network. Finally, the distributed tracking and consensus in networked multi-agent systems problem shows that information sharing among users improves the quality of local estimates and helps avoid conflicting and inefficient distributed decisions. It also reveals the effect of the graph topologies and incomplete node measurements on the speed of achieving distributed decision and final consensus accuracy.

Contents

List of Figures

Glossary

$\|G\|_\infty$	H-infinity norm of matrix G. $\|G\|_\infty = \sup_\omega \overline{\sigma}\big[G(j\omega)\big]$,where $\overline{\sigma}$ is the maximum singular value of the matrix $G(j\omega)$.
ω	Angular frequency: $\omega = 2\pi f$, where f is the ordinary frequency.
$A \otimes B$	Kronecker product of matrix A and B.
diag$(d_1, \cdots, d_N)$	A diagonal matrix with $(d_1, \cdots, d_N)$ on its main diagonal.
$\lambda(G)$	Eigenvalue of matrix G.
$A \oplus B$	Direct sum of matrix A and B.
$\mathcal{L}\{x(t)\}$	Laplace transform of a continuous signal $x(t)$, $X(s) = \mathcal{L}\{x(t)\}$.

Chapter 1

Introduction

1.1 Background and Motivation

With dramatic developments in wireless communication technology and network science, the restrictions on information exchange between spatially distributed agents have been relaxed and traditional control systems have been extended in a distributed manner. The application of distributed control is made possible to large-scale control objects and physically distributed agents. *Networked multi-agent system* is a system, in which multiple agents with distributed control architecture exchange information and interact with each other in the form of data packages through a network.

One unique feature of networked multi-agent system is that the agents are interconnected but physically decoupled, whereas are coupled via information exchange and interaction with each other when trying to accomplish certain assigned tasks. Besides, with information exchange among agents, distributed control laws are generated according to not only the local feedback, but also the messages from other agents. In networked multi-agent systems, complicated tasks, which are too difficult for single agent, may be accomplished through collaboration of agents over network. Networked multi-agent system also shows advantages in simple structure, low cost, robustness,

scalability, and enhanced functionality.

In recent years, cooperation and coordinated control in networked multi-agent systems has received significant attention from scientific community. Typical application of multi-agent systems can be found in smart grid [1–4], automated highway systems [5–9], air traffic control [10, 11], congestion control on the internet [12, 13], and sensor networks [14–16] and so forth. In those applications, an agent can represent a power plant, an airplane, a vehicle, an internet router or even a smart sensor with computation capability.

However, the applications of networked multi-agent systems also bring in a couple of issues that need to be addressed.

Compared to single agent case, the dynamics of multi-agent systems are more complicated and more difficult to predict due to interactions among agents. Some problems could be very complicated under certain condition, such as multi-agent markov game [17].

Communication links in networked multi-agent systems are not perfect and unpredictable, which introduces time delays and packet losses. In wireless network, information is transmitted over a wireless highly constrained communication medium. The signal propagation through wireless channels experiences random power fluctuations over time, which is multipath interference caused by reflection and attenuation from objects. These power fluctuations cause time-varying data rates and intermittent connectivity, which inevitably introduces time delays and packet losses.

The network topology of agents greatly affects the stability and performance of networked multi-agent systems. In networked multi-agent systems the assigned task is cooperatively accomplished by group of agents. The performance issue depends on not just the agent dynamics but also the agent interaction over network.

The task assignment and time synchronization among agents over network are issues need to be considered. Task decomposition and assignment are traditionally solved in a centralized manner. However, for distributed agents over network, there is no centralized controller to conduct such task assignment. Also, asynchronization issue will rise among agents which communicate over

network with time delays and packet losses.

To address and overcome these technological challenges in networked multi-agent systems, inter-disciplinary designs and strategies should be used from areas such as control theory, communication theory, computation. It is very clear that control theory can definitely benefit from other closely related disciplines, such as algebraic graph theory, communication theory, distributed computation and network information theory [18].

In this report, our work on information theory and cooperative control in networked multi-agent systems is presented along with short survey of recent work in the following topics: information theory for tracking in networked control systems, distributed demand-side management in smart grid with communication constraints, disturbance attenuation and string stabilization in automated highway systems, and distributed tracking and consensus seeking in wireless sensor networks.

1.2 Literature Review

1.2.1 Infomtion Theory in Networked Control Systems

In networked control systems, agents or system components, which may be distributed over geographically disparate locations, coordinate with each other by exchange information through a communication network [19–21,32]. Since cooperation and communication is over a network with certain uncertainties, such as time delay, noisy communication links, changing network topology. Analyzing the performance of such system not only depends on the agent dynamics and coordination, but also relies on the condition of the network. In literature, network information theory has provided useful tools in analysis of such system and its application has become an important topic in this area.

There has been considerable attention given to the application of information theory in the stabilization of networked control systems [23–40]. Previous work in [23–25] has shown that

stabilization of a linear and time-invariant plant, requires that the channel capacity C^{Cap} to be larger than $\sum_i \max\{0, \log_2(|\lambda_i(A)|)\}$, where the sum is over all unstable eigenvalues of the dynamic matrix of the state-space representation of the plant. The papers [26, 28–32] have shown that the extra rate $C^{\text{Cap}} - \sum_i \max\{0, \log_2(|\lambda_i(A)|)\}$ is critical for performance, as measured by the expected power of the state of the plant. The authors in [29,30,41] studied fundamental limitations in disturbance rejection in feedback systems and extended the Bode's integral equation for the case where the preview is made available to the controller via a general, finite capacity, communication channel. Authors in [27,28,33,34] provided necessary conditions on the data rate of each channel for stabilization of a linear and time-invariant plant over a packet-based network. In [42–44], authors introduced ϵ-entropy to measure the variety of a linear continuous time system, which is defined as the difference between ϵ-entropy rates of system input and output. Relations within variety, Bode sensitivity integral and H_∞entropy were then derived.

Later, Authors in [35,49–51] extended the Shannon channel capacity which is not sufficient to characterize a noisy communication channel in feedback loop to stabilize an unstable scalar linear system. They introduced "Anytime" channel capacity and showed its necessity and sufficiency in characterizing the stabilization of networked control systems over noisy communication channels.

Further, the authors in [36–40] extended the problem of feedback stabilization of a linear time invariant (LTI) plant with a feedback communication channel with finite capacity to the one with channel that is subject to a constraint on the channel signal-to-noise ratio (SNR). The authors obtained a closed-form expression for the infimal SNR for stabilizability subject to quantization error of a memoryless AWGN channel.

On the other hand, other recent work has been focused on tracking issues in networked feedback systems. The work in [52–54] shows that a necessary condition for efficient tracking is that the information flow from the reference signal to the output should be greater than the information flow between the disturbance and the output. The concept of directed information and mutual information were introduced to provide necessary conditions for tracking in linear feedback systems. Meanwhile, the work in [56–58] defines conditions for tracking such that tracking error has finite

energy. Following the same approach, in this chapter we find conditions for tracking such that the power of the tracking error stays finite. The authors in [57, 58] obtained information theoretical conditions for tracking in linear time-invariant control systems, where the closed loop contains a channel in the feedback loop. The authors provided an upper bound for the mutual information rate between the feedback signal and the reference input signal and showed that this rate must be maximized to improve the tracking performance.

1.2.2 Distributed Demand-side Management over Smart Grid

New solutions for electrical energy generation have been proliferated in recent years, including wind and solar power, most of which are intrinsically distributed and inherently unpredictable. The integration of such distributed energy resources (DER), mainly of renewables, into power generation mix has brought about the concept of *Smart Grid*. In essence, the smart grid is a power network composed of intelligent nodes that can operate, communicate, and interact autonomously to efficiently deliver power and electricity to their consumers [59]. With the help of two-way communication, distributed optimization and control technologies, the smart grid is expected to be capable of integrating DERs and loads more reliably, efficiently and sustainably [60].

The concept of a *microgrid* plays an important role as the building block of smart grid due to its flexibility and scalability [61]. A microgrid is a localized grouping of distributed generation (DG) (i.e. fuel cells, photovoltaic array, wind power, etc), distributed storage (i.e., flywheels, batteries) and controllable loads. A microgrid can operate both in grid-connected and islanded mode [62]: in grid-connected mode, the microgrid is connected to a traditional power grid as usual. When a fault occurs in the upstream grid, however, it can be disconnected from the grid and function autonomously in an islanded mode. The microgrid aids future smart grid for its capability of aggregating DER and presenting itself to the electrical network as a controllable entity [63].

Demand-side management (DSM) is an essential aspect of current and future smart grid system that allows the utility companies and consumers (e.g. residences and businesses) to interact with

each other in order to achieve objectives such as energy efficiency and demand profile improvement [64, 65], utility and cost optimization [69] and emission control [64].

There has been considerable attention on demand-side management in smart grids, especially within microgrids [64–67, 69–71]. Optimization techniques such as dynamic, stochastic and robust programming are widely used as tools in demand-side management for microgrids [62]. For example, the authors in [64] designed a three-step control and optimization strategy and focused on the control algorithms to reshape the energy demand profile of a large group of buildings and their requirements on the smart grid. An online learning algorithm for residential demand response to reduce residential energy costs and smooth energy usage was proposed in [65]. A stochastic program was formulated by applying microgrid technology to minimize the overall cost of electricity and natural gas for a building operation in [66]. The authors in [67] proposed a robust optimization model to adjust the hourly load level of a given consumer in response to hourly electricity prices.

The above approaches have often formulated a system-level optimization problem based on at centralized objective function, which requires collective information at a centralized controller. On the other hand, distributed analytical techniques, such as game theory and distributed control, are ideal in solving problems through coordination and interaction among customers. Under the non-cooperative user assumption, the authors in [68] developed a distributed auctioning game to coordinate the interaction among all customers to meet an optimal aggregate load profile. The authors in [69] developed a demand-side energy consumption scheduling game to optimize the aggregate load of the users by coordination of user energy usage through a two-way digital communication infrastructure. The authors in [70] applied a distributed subgradient method in a demand response problem to minimize the total cost of electricity from all residences and the total user dissatisfaction, subject to the individual load consumption constraints. In [71], the authors formulated an optimal power scheduling problem in a smart grid subject to generation, consumption and storage constraints and developed a decentralized average message passing method through message exchange within a neighborhood and parallel optimization.

This paper considers a framework in which distributed customers equipped with renewable dis-

tributed generation (RDG) interact and trade energy within a mircogrid. We consider distributed demand-side management for such a microgrid by optimizing the aggregate load of users through coordination among them. Compared to the aforementioned distributed approaches [68–71], we take advantage of cooperation among users in energy consumption coordination and apply distributed control to achieve aggregate load optimization. We first propose a distributed networked control strategy based on information exchange among nodes to optimize the aggregate load of users and maintain supply-demand balance. Next, by taking into account the possible delay in communication between users and management system within microgrids, we propose a distributed networked control strategy with prediction to solve the aggregate load optimization problem with communication delays.

1.2.3 String Stablization in Automated Highway Systems

Automated highway system (AHS) [6, 7, 72–75] is a vehicle platoon among which each vehicle maintains a fixed distance behind the preceding one by using closed loop control. AHS is designed to solve traffic congestion problem and improve the traffic management in transportation system. In this vehicle platoon, vehicles are coupled with each other through feedback loops and the disturbance in one vehicle could propagate and affect other vehicles in the string. Adaptive cruise control (ACC) [76,77,79] has been proposed as an enhancement over existing cruise controllers on ground vehicles. Adaptive cruise control (ACC) systems control the vehicle speed to follow a drivers set speed closely when no lead vehicle is insight. When a slower leading vehicle is present in an AHS, the controlled vehicle will follow the lead vehicle at a safe distance. AHS and ACC research first began in the 1960s [80], and has received ever-growing attention in the last decade. However, their commercial implementation is not possible until recently with significant progresses in sensors, actuators, and other enabling technologies.

Over the last decade, many different approaches have been proposed for the design of algorithms [8, 9, 79, 81–89]. In the earlier work, the focus has been on the performance of the host

vehicle. The performance was usually evaluated based on 2- car platoons. The effect of AHS on a string of vehicles has not received much attention. The term *string stability* [5] is introduced to describe the amplification along the string of the response to a disturbance to the lead vehicle. It is widely known that when the transfer function from the range error of a vehicle to that of its following vehicle has a magnitude less than 1, string stability is guaranteed [76]. [6, 72] developed a formal definition of string stability and used the norm induced by the signal norm to characterize it. The string stability ensures that range errors decrease as they propagate along the vehicle stream. In a long string of vehicle with tight formation, string stability could be not ensured when each vehicle uses only the relate spacing information of its predecessor.

To achieve string stability with constant inter-vehicle spacing, researchers tried to design different control schemes and vehicle-to-vehicle communication was shown to be necessary [90]. Yanakiev and Kanellakopoulos used a simple spring-mass-damper system to demonstrate the idea of string stability and show the string stability criterion for constant time headway and variable time-headway policies [73]. Swaroop and Hedrick proved [6, 7, 72], among many other interesting things that if the coupling between two vehicles is weak enough, the controlled system is string stable. In [7, 9] authors proposed "multi-look ahead"control scheme, in which controller on each vehicle utilizes only information about itself and the one directly ahead. [8] proposed predecessor and leader following control scheme in which each vehicle has access to information from the lead vehicle. [89] presented distributed receding horizon control algorithms and derived sufficient conditions that guarantee asymptotic stability, leader-follower string stability, and predecessor-follower string stability, following a step speed change in the platoon.

Despite predecessor-following and predecessor and leader following control strategies, the unidirectional scheme was proposed in which each vehicle not only communicates with its predecessor but also its next vehicle in the string. [84–88] considered the string stability with the unidirectional nearest neighbor communications. The authors introduced "time headway policy" and provided conditions under which the L_2-norm and the L_∞-norm of states are bounded and derived a formula for the minimal time headways required to guarantee string stability.

Communication delay among vehicles would worsen the string stability by increasing the possible actuation time of each vehicle. Authors in [74,75,77,78] studied the effect of communication delay in string stability of vehicle platooning and showed that string stability could be guaranteed when the constant time headway is much larger that the time delay.

Although string stability among vehicles with identical controllers has been extensively explored, heterogeneous string stability received less attention. [81–83] considered string stability in vehicle platoon with heterogeneous controllers and constructed a controller design procedure that gives string stability and robustness to external disturbances for heterogeneous vehicle strings.

1.2.4 Tracking and Consensus over Sensor Networks

Multi-sensor tracking problems have attracted the attention of many researchers in robotics, systems, and control theory over the past three decades [91–94]. Target tracking problems are of great importance in surveillance, security, and information systems for monitoring the behavior of agents using sensor networks, such as tracking pallets in warehouses, vehicles on roadways, or firefighters in burning buildings [95,96]. With the introduction of the concept of consensus, distributed tracking and coordination without any fusion center has also received considerable attention in recent years [98,99].

Distributed consensus estimation in sensor networks has been studied with both fixed as well as time varying communication topologies, taking into account issues such as link failure, packet losses and quantization or additive channel noise [14,99,101–111]. Olfati-Saber and Murray [101] considered the average consensus for first-order integrator networks with fixed and switching topologies. Kingston and Beard [102] extended the results of [101] to the discrete-time models and relaxed the graph condition to instantaneously balanced, connected-over-time networks. Xiao and Boyd [14] considered discrete-time distributed averaging consensus over fixed and undirected graphs. They designed the weighted adjacency matrix to optimize the convergence rate by semi-definite programming. Huang and Manton [108] considered the discrete-time average consensus

with fixed topologies and measurement noises. They introduced decreasing step size in the protocol to attenuate the noises. Li and Zhang [109–111] considered the continuous-time average consensus with time-varying topologies and communication noises, where time-varying consensus gains are adopted. They gave a necessary and sufficient condition for mean square average-consensus with fixed graph topologies and sufficient conditions for mean square average-consensus and almost sure consensus with time-varying graph topologies.

On the other hand, the distributed consensus tracking over networks with and without noiseless communication links among nodes have also been considered [113–118]. Recent work in [113, 114] considered the distributed consensus tracking over a fixed graph with noiseless communication among nodes. A distributed Kalman filter with embedded consensus filters was proposed in [113] and further extended to heterogeneous and nonlinear sensing models in [114]. Distributed Kalman filtering using one-step weighted averaging was considered in [115]. Each node desires an estimate of the observed system and communicates its local estimate with neighbors in the network. Then new estimate is formed as a weighted average of the neighboring estimates, where the weights are optimized to yield a small estimation error covariance. In [116], the authors presented a distributed Kalman filter to estimate the state of a sparsely connected, large-scale, dynamical system. The complex large-scale systems is decomposed spatially into low-order overlapping subsystems. A fusion algorithm using bipartite fusion graphs and local average consensus algorithms is applied to fuse observations for those overlapping subsystems. A tracking control problem for multi-agent consensus with an active leader and variable interconnection topology was considered in [117], where the state of the considered leader keeps changing and may not be measured. A neighbor-based local controller together with a neighbor-based state-estimation rule is given for each agent to track such a leader. In [118], the authors proposed a greedy stepsize sequence design to guarantee the convergence of distributed estimation consensus over a network with noisy links.

Distributed tracking with consensus, addressed in this chapter and previous work [121, 122], refers to the problem that a group of nodes need to achieve an agreement over the state of a dynamical system by exchanging tracking estimates over a network. For instance, space-object tracking

with a satellite surveillance network which consists of fixed nodes that are connected together and mobile nodes that could only have active links with other nodes within their communication radius could benefit from such distributed tracking with consensus, due to the fact that individual sensor nodes may not have enough observations of sufficient quality [119]. Thus, different sensor nodes may arrive at different local estimates regarding the same space object of interest [119]. Information exchange among nodes may improve the quality of local estimates and consensus estimates may help avoid conflicting and inefficient distributed decisions. Other applications of this problem include flocking and formation control, real-time monitoring, target tracking and global positioning system (GPS) [119, 120]. In [122], the performance analysis of distributed tracking with consensus on noisy time-varying graphs was addressed and later the algorithm of distributed tracking with consensus with incomplete data was proposed without theoretical proof in [121] .

1.3 Outline of This Dissertation

The research has several threads and this document is organized by focusing on a different topic in each chapter as follows.

Chapter 2 considers the problem of tracking in networked control systems – leader-follower systems under communication constraints. By using information theory and control theory, we provide necessary information theoretic conditions on the channel data rate of each communication link for tracking in such system.

Then, Chapter 3 focuses on distributed demand-side management in smart grid, which is a networked multi-agent system with RDG, loads and DS. We propose a distributed control based energy scheduling method by optimizing the aggregate load of users with RDG through user energy usage coordination to maintain load demand balance.

Chapter 4 analyzes the string stability problem in automated highway systems (AHS) with an inter-vehicle communication network. We consider controlling a platoon of vehicles in which

each vehicle tries to maintain a fixed distance from its predecessor, which is an instance of the "string stability" problem. We propose a control protocol – leader-following consensus protocol and derive the sufficient conditions in terms of communication topology and control parameters, for string stability in such system.

Later, in Chapter 5 we consider the problem of distributed tracking in wireless sensor networks (WSN) through cooperative control. We formulates the problem of the distributed tracking in a network of sensors with a time-varying network graph, incomplete node data and noisy communication links. A distributed tracking with consensus algorithm is proposed and the convergence of the proposed algorithm and the steady state behavior are presented.

Finally, Chapter 6 draws conclusions from the completed work and outlines future research work.

Chapter 2

Information Theory for Tracking in Networked Control Systems

2.1 Introduction

In networked control system problems, understanding the fundamental relationship between how the control parts and the communication parts of the distributed system interact is significant for controller and communication channel design. Previous work in [23, 24] has shown that stabilization of a linear and time-invariant plant, requires that the data rate C to be larger than $\sum_i \max\{0, \log_2(|\lambda_i(A)|)\}$, where the sum is over all unstable eigenvalues of the dynamic matrix of the state-space representation of the plant. The papers [26, 29] have shown that the extra rate $C - \sum_i \max\{0, \log_2(|\lambda_i(A)|)\}$ is critical for performance, as measured by the expected power of the state of the plant.

On the other hand, recent work has been focused on tracking issues in networked feedback systems. The work in [41] shows that a necessary condition for efficient tracking is that the information flow from the reference signal to the output should be greater than the information flow between the disturbance and the output. Meanwhile, the work in [57] defines conditions for track-

ing such that tracking error has finite energy. Following the same approach, in this chapter we find conditions for tracking such that the power of the tracking error stays finite. The authors in [57] obtained information theoretical conditions for tracking in linear time-invariant control systems, where the closed loop contains a channel in the feedback loop. The authors provided an upper bound for the mutual information rate between the feedback signal and the reference input signal and showed that this rate must be maximized to improve the tracking performance.

In this chapter, we introduce a general framework for tracking in leader-follower systems under communication constraints, where the leader system, follower system and the corresponding controllers are spatially distributed and connected over communication links. The communication channels are used to exchange information and control signals among spatially distributed system components. We consider the particular case in which both the forward link from the reference signal input and the feedback link from the system output contain communication channels with finite data rate.

For this particular problem, we derive necessary conditions on channel data rate of the forward and feedback links for tracking in the leader-follower systems. Then, we show the effect of existence of the feedback link on the required channel data rate of the forward link, when the feedback link is noisy. By considering the forward and feedback channels as one cascade channel, we also provide a lower bound for the data rate of the cascade channel for the system to track the reference signal such that the tracking error has finite second moment.

The chapter is divided into five sections as follows: Section 2.2 introduces the notation and the main definitions and properties from information theory. The problem formulation of leader-follower systems under communication constraints is given in Section 2.3, where we describe the assumptions on communication channels, dynamic systems and the reference signal. In Section 2.4, we show the necessary conditions on individual channels for tracking in leader-follower system and provide a lower bound on the data rate of the cascade channel by considering the forward channel and feedback channel together. In Section 2.5, these above results are extended to the case that leader and follower systems have different system models. In Section VI, we study special

cases and demonstrate our results in Section 2.6. The conclusions are provided in Section 2.7.

2.2 Definitions and Properties

In the following, we present the definitions and properties used later in this chapter.

Definition 2.2.1. *(Entropy): For a given discrete random variable* $\mathbf{x}$*, the entropy is defined by:*

$$h(\mathbf{x}) = \sum_{\mathbf{x}} p(\mathbf{x}) \log p(\mathbf{x}),$$

where $p(\mathbf{x})$ *is the probability density function of* $\mathbf{x}$*.*

Definition 2.2.2. *(Mutual Information): The mutual information between discrete random variables* $\mathbf{x}$ *and* $\mathbf{y}$ *is defined as*

$$I(\mathbf{x};\mathbf{y}) = \sum_{\mathbf{x}\in\mathcal{A}} \sum_{\mathbf{y}\in\mathcal{B}} p(\mathbf{x},\mathbf{y}) \log_2 \frac{p(\mathbf{x},\mathbf{y})}{p(\mathbf{x})p(\mathbf{y})},$$

where $p(\mathbf{x},\mathbf{y})$ *is the joint probability density function of* $\mathbf{x}$ *and* $\mathbf{y}$*.*

Definition 2.2.3. *(Entropy Rate): For a given stochastic process* $\mathbf{a}$*, the entropy rate is defined as [128]:*

$$h_\infty(\mathbf{a}) = \limsup_{k\to\infty} \frac{h(\mathbf{a}^k)}{k}.$$

Definition 2.2.4. *(Information Rate): Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be stochastic processes. The mutual information rates are defined as [29]:*

$$I_\infty(\mathbf{a};\mathbf{b}) = \limsup_{k\to\infty} \frac{I(\mathbf{a}^k;\mathbf{b}^k)}{k}.$$

where $I(\mathbf{a}^k;\mathbf{b}^k)$ *is mutual information between* $\mathbf{a}^k$ *and* $\mathbf{b}^k$ *and it can be obtained as follows:*

$$I(\mathbf{a}^k;\mathbf{b}^k) = h(\mathbf{a}^k) - h(\mathbf{a}^k|\mathbf{b}^k). \tag{2.1}$$

Definition 2.2.5. *(Directed Mutual Information and Directed Information Rate): Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be stochastic processes. The directed mutual information is defined as follows [29]:*

$$I(\mathbf{a}^k \to \mathbf{b}^k) = \sum_{i=1}^{k} I(\mathbf{a}^i; \mathbf{b}(i)|\mathbf{b}^{i-1}),$$

and the directed information rate is given by

$$I_\infty(\mathbf{a} \to \mathbf{b}) = \limsup_{k\to\infty} \frac{I(\mathbf{a}^k \to \mathbf{b}^k)}{k}.$$

Definition 2.2.6. *(Channel Capacity): For channel* CH_i *with input* $\mathbf{x}_i$ *let the corresponding output be denoted by* $\hat{\mathbf{x}}_i$*, define the error function* $E_i(k)$ *at time step* k *as* $E_i(k) = \begin{cases} 1 & \mathbf{x}_i \neq \hat{\mathbf{x}}_i \\ 0 & \mathbf{x}_i = \hat{\mathbf{x}}_i \end{cases}$*. The channel capacity* C_i^{Cap} *is defined as the supremum of all achievable rates,*

$$C_i^{Cap} = \sup_{p(\mathbf{x}_i)} I(\mathbf{x}_i; \hat{\mathbf{x}}_i). \tag{2.2}$$

Properties 2.2.7. *Assume that* $\mathbf{a}$*,* $\mathbf{b}$*,* $\mathbf{c}$*,* $\mathbf{d} \in \mathbb{R}^n$ *are random variables, and* f *and* g *are real functions. All the following properties may be found in [128, 133].*

(P1) $I(\mathbf{a}; \mathbf{b}) = I(\mathbf{b}; \mathbf{a}) \geq 0$ *and* $I(\mathbf{a}; \mathbf{b}|\mathbf{c}) = I(\mathbf{b}; \mathbf{a}|\mathbf{c}) \geq 0$*.*

(P2) $I((\mathbf{a}, \mathbf{b}); \mathbf{c}|\mathbf{d}) = I(\mathbf{b}; \mathbf{c}|\mathbf{d}) + I(\mathbf{a}; \mathbf{c}|(\mathbf{b}, \mathbf{d}))$*.*

(P3) *If* f *and* g *are measurable functions then* $I(f(\mathbf{a}); g(\mathbf{b})|\mathbf{c}) \leq I(\mathbf{a}; \mathbf{b}|\mathbf{c})$ *and equality holds if* f *and* g *are invertible.*

(P4) *Given a function* $f : \mathcal{C} \to \mathcal{C}'$*, it follows that* $I(\mathbf{a}; f(\mathbf{c})|\mathbf{c}) = 0$*.*

(P5) $h(\mathbf{a}|\mathbf{b}) = h(\mathbf{a} - g(\mathbf{b})|\mathbf{b})$*.*

(P6) $h(\mathbf{a}|\mathbf{b}) \leq h(\mathbf{a})$ *with equality if* $\mathbf{a}$ *and* $\mathbf{b}$ *are independent.*

(P7) *Let* $\mathbf{a} \in \mathbb{R}^n$ *have mean* $\boldsymbol{\mu}$ *and covariance Cov*$\{\mathbf{a}\}$. *Then*

$$h(\mathbf{a}) \leq \frac{1}{2}\log_2\left((2\pi e)^n \det\big(Cov\{\mathbf{a}\}\big)\right)$$

with equality if $\mathbf{a}$ *has a multivariate normal distribution.*

(P8) $I((\mathbf{a},\mathbf{b});\mathbf{c}) = I(\mathbf{a};\mathbf{c}|\mathbf{b}) + I(\mathbf{b};\mathbf{c})$.

(P9) (Fano's inequality) *For channel* CH_i *with input* $\mathbf{x}_i$ *and corresponding output* $\hat{\mathbf{x}}_i$. *Let the probability of error as* $P_{e,i} = Pr\{\hat{\mathbf{x}}_i \neq \mathbf{x}_i\}$. *Then*

$$h(P_{e,i}) + P_{e,i}\log_2|\mathcal{X}_i| \geq h(\mathbf{x}_i|\hat{\mathbf{x}}_i), \tag{2.3}$$

where $\mathcal{X}_i$ *is the alphabet for input* $\mathbf{x}_i$.

(P10) $I((\mathbf{a},\mathbf{b});\mathbf{c}) \geq I(\mathbf{b};\mathbf{c})$.

2.3 Information Theory in Leader-Follower Systems

Consider the following networked control system as in Fig. 2.1. There are two physical systems P_1 and P_2 controlled by corresponding controllers $\mathcal{C}_1$ and $\mathcal{C}_2$ over communication channels CH_i for $i \in \{11, 12, 2, 3, 4\}$ with finite rates. The systems and controllers are spatially distributed and connected over communication links. The communication channels are used to exchange information and control signals among spatially distributed system components such as controllers, actuators and sensors.

In our framework, the two linear and time-invariant systems denoted as the leader P_1 and the follower P_2 are assumed to be identical. The framework could be further extended to more general leader-follower network with multiple leaders and followers interconnected over communication links. There are many possible applications of such a general framework including, for example,

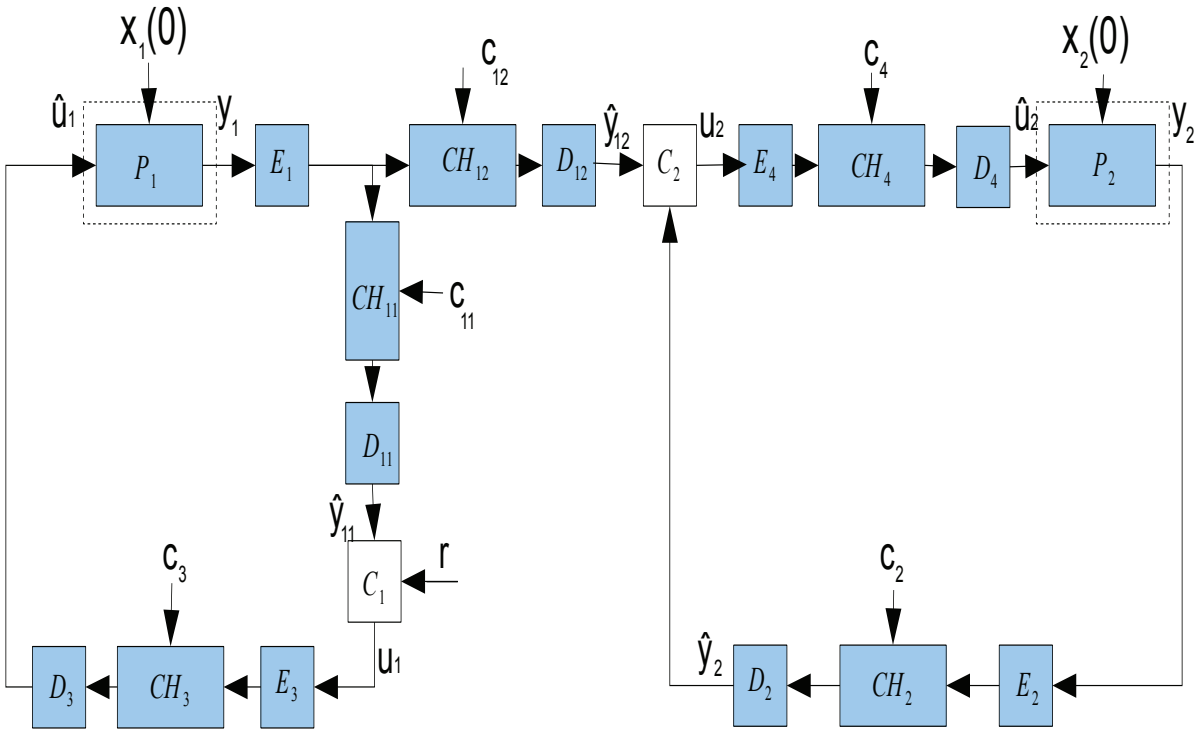


Figure 2.1: A general model of a networked control system with two plants.

distributed control of power plants in smart-grid, where control input could be applied to one generator and other generators could act as followers to track the state of that generator. Another possible application is the formation control of homogeneous robots with exterior control applied only to the leader [125–127]. However, such general network topologies are beyond the scope of this report.

Suppose that the reference signal $\mathbf{r}(k)$ is only available for the leader system P_1. The follower system P_2 does not have information of $\mathbf{r}(k)$ and has to track the state of the leader system based only on the output of the plant P_1 over a communication network. The goal is to find the lower bound of channel data rate for channel CH_i to convey enough information to both controllers $\mathcal{C}_1$ and $\mathcal{C}_2$ to generate efficient feedback control signals such that the plants P_1 and P_2 could track $\mathbf{r}(k)$ accurately. Note that, the reference signal $\mathbf{r}(k)$ may not be available for follower system P_2 as a

result of the high cost of information delivery to each plant due to long-range spatial separation between systems or a large number of follower systems (not considered here).

In the following, we formulate the discrete-time state-space representation for the leader and follower systems[1]:

$$\begin{aligned} \mathbf{x}_j(k+1) &= F\mathbf{x}_j(k) + G\hat{\mathbf{u}}_j(k), \\ \mathbf{y}_j(k) &= H\mathbf{x}_j(k), \qquad k \geq 0, j = 1, 2, \end{aligned} \tag{2.4}$$

where the states $\mathbf{x}_j(k)$ takes values in $\mathbb{R}^n$ and the received control input $\hat{\mathbf{u}}_j(k)$ takes values in $\mathbb{R}^r$. The initial state $\mathbf{x}_j(0)$ is a zero mean Gaussian random variable with covariance matrix Σ_{0j}. The state is observed by a sensor that generates the measurement $\mathbf{y}_j(k)$ taking values in $\mathbb{R}^q$.

We assume that the pair (F, G) is controllable and the pair (F, H) is observable based on the fact that follower system P_2 tries to track the state of leader system P_1, which requires system state $\mathbf{x}_1$ to be observable and $\mathbf{x}_2$ to be controllable. Since the pair (F, H) is observable, the system state $\mathbf{x}_j$ could be sufficiently determined by output signal $\mathbf{y}_j$. In order to simplify the derivation in next section and achieve theoretical results, we assume that system state could be estimated by $\mathbf{x}_j = L\mathbf{y}_j$ for $j = 1, 2$[2], where matrix $L \in \mathbb{R}^{n\times q}$ is a transformation matrix.

Here we define the tracking errors of systems P_1 and P_2 as $\boldsymbol{\xi}_1(k) = \mathbf{r}(k) - \mathbf{x}_1(k)$ and $\boldsymbol{\xi}_2(k) = \mathbf{x}_1(k) - \mathbf{x}_2(k)$, where $\boldsymbol{\xi}_j(k)$ is a stochastic process with mean $\boldsymbol{\mu}_j$ and covariance $\text{Cov}\{\boldsymbol{\xi}_j\}$, for $j = 1, 2$. Assume that the matrix $F = \begin{bmatrix} F_s & 0 \\ 0 & F_u \end{bmatrix}$ where $0 < |\lambda_i(F_s)| < 1$ and $|\lambda_i(F_u)| \geq 1$. Therefore, F^k is invertible $\forall k$.

The encoders and decoders are described as follows:

1) Encoder: At every time step k, encoder ϵ_i calculates and transmits the vector $s_i(k)$ for

[1]For ease of mathematical derivation, we only consider identical model systems.

[2]The state-output relation assumption may limit the use of this chapter's results in some practical applications. We are currently working on relaxing this assumption on further extensions.

$i = 1, \cdots, 4$, according to the following functional structure:

$$\mathbf{s}_1(k) = \epsilon_1(\mathbf{y}_1^k),\ \mathbf{s}_2(k) = \epsilon_2(\mathbf{y}_2^k),$$
$$\mathbf{s}_3(k) = \epsilon_3(\mathbf{u}_1^k),\ \mathbf{s}_4(k) = \epsilon_4(\mathbf{u}_2^k),$$

where $\mathbf{s}_i(k)$ takes values in $\mathbb{R}^m$ and $\mathbf{y}_i^k = \{\mathbf{y}_i(1), \cdots, \mathbf{y}_i(k)\}$.

2) Discrete-time memory-less channel (DMC): Let $\mathfrak{S}_i$ and $\mathfrak{Z}_i$ be given input and output alphabets, along with a white stochastic process, denoted as $\mathbf{c}_i$, with alphabet $\mathfrak{C}_i$. Consider the mapping $\mathcal{F}_i : \mathfrak{S}_i \times \mathfrak{C}_i \to \mathfrak{Z}_i$ for $i \in \{11, 12, 2, 3, 4\}$ such that the following maps: $\mathbf{z}_i(k) = \mathcal{F}_i(\mathbf{s}_i(k), \mathbf{c}_i(k))$, where $\mathbf{c}_i$ is the channel noise.

3) Decoder: We consider the decoder for channel CH_i is of the following form:

$$\hat{\mathbf{y}}_{11}(k) = D_{11}^k(\hat{\mathbf{y}}_{11}^{k-1}, \mathbf{z}_{11}^k),\ \hat{\mathbf{y}}_{12}(k) = D_{12}^k(\hat{\mathbf{y}}_{12}^{k-1}, \mathbf{z}_{12}^k),$$
$$\hat{\mathbf{y}}_2(k) = D_2^k(\hat{\mathbf{y}}_2^{k-1}, \mathbf{z}_2^k),\ \hat{\mathbf{u}}_1(k) = D_3^k(\hat{\mathbf{u}}_1^{k-1}, \mathbf{z}_3^k),$$
$$\hat{\mathbf{u}}_2(k) = D_4^k(\hat{\mathbf{u}}_2^{k-1}, \mathbf{z}_4^k).$$

The controllers $\mathcal{C}_1$ and $\mathcal{C}_2$ are defined as follows:

$$\begin{aligned} \mathcal{C}_1 :&\quad \mathbf{u}_1(k) = f_1\big(\mathbf{e}_1^k\big) \text{ with } \mathbf{e}_1(k) = H\mathbf{r}(k) - \hat{\mathbf{y}}_{11}(k), \\ \mathcal{C}_2 :&\quad \mathbf{u}_2(k) = f_2\big(\mathbf{e}_2^k\big) \text{ with } \mathbf{e}_2(k) = \hat{\mathbf{y}}_{12}(k) - \hat{\mathbf{y}}_2(k), \end{aligned} \tag{2.5}$$

where we assume that reference signal $\mathbf{r}(k)$ has finite power such that $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$.

2.4 Information Theoretic Conditions for Tracking

In order to be able to derive useful results on channel capacity and conditions for optimal control, first we simplify the general model in Fig. 2.1 as shown in Fig. 2.2. In Fig. 2.2, the controllers are assumed to be directly connected to actuators that operate the systems, so that we can assume that the channels CH_3 and CH_4 are lossless with no delays.

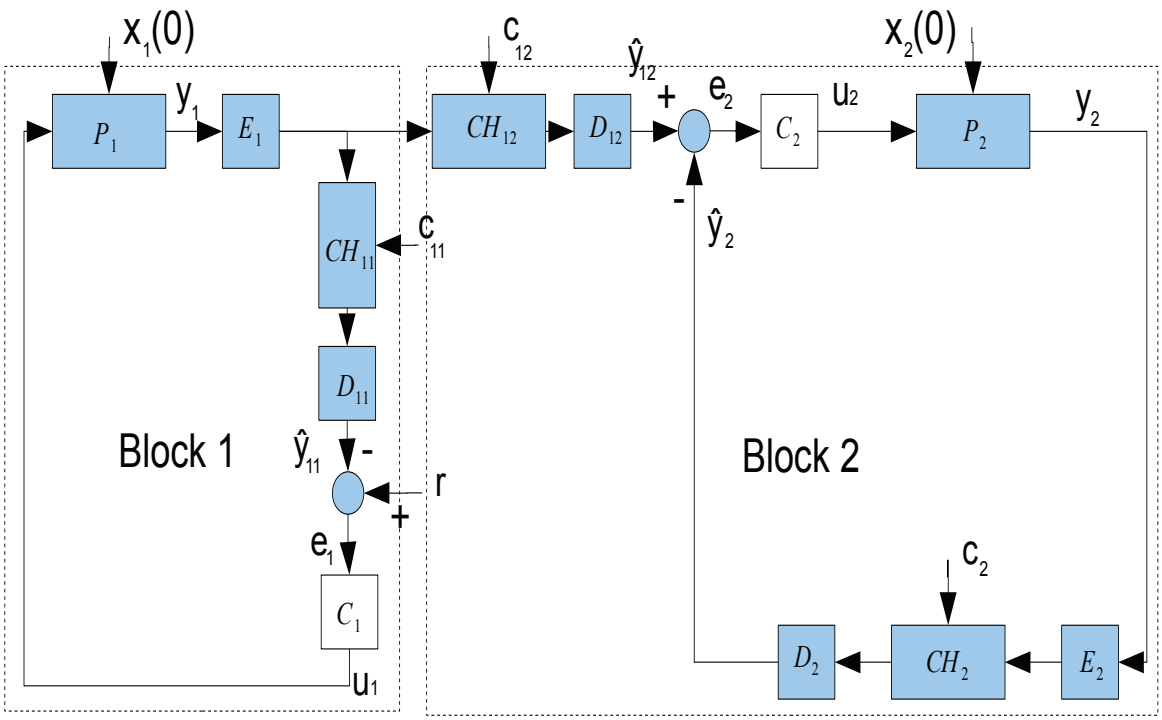


Figure 2.2: A simplified model of a networked control system with two plants.

2.4.1 Necessary Conditions on Individual Channel

Channel CH$_{11}$

Consider block 1 in Fig. 2.2. Plant P_1 is not affected by the second plant P_2. The block is a closed-loop system with communication channel in feedback link as in [57]. Before proceeding with our results, we extend Lemma 2 in [57] without assuming that $H = I$ and state it as Lemma 2.4.1.

Lemma 2.4.1. *Consider the closed-loop system in block 1 in Fig. 2.2, where plant P_1 is an LTI system described by (2.4). Assume that the pair (F, G) is controllable and (F, H) is observable.*

Assume $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$ *for reference signal* $\mathbf{r}(k)$. *If* $\mathbb{E}[\boldsymbol{\xi}_1(k)^T\boldsymbol{\xi}_1(k)] < \infty$, *then*

$$\lim_{k\to\infty} \frac{I(\mathbf{x}_1(0); \mathbf{e}_1^k|\mathbf{r}^k)}{k} \geq \lim_{k\to\infty} \frac{I(\mathbf{x}_1(0); \mathbf{e}_1^k)}{k} \geq \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}, \tag{2.6}$$

where $\mathbf{e}_1^k = H\mathbf{r}^k - \hat{\mathbf{y}}_{11}^k$.

Proof. Note that the matrix F can be written in the form $F = \begin{bmatrix} F_s & 0 \\ 0 & F_u \end{bmatrix}$ where F_s corresponds to the stable subspace ($0 < |\lambda_i(F_s)| < 1$) and F_u corresponds to the marginally stable and unstable subspace ($|\lambda_i(F_u)| \geq 1$). If $F = F_s$, from (P1) we just have $I(\mathbf{x}_1(0); \mathbf{e}_1^k) \geq 0$. For any control sequence, the system remains stable. Hence, without loss of generality, we can restrict our attention to matrix $F = F_u$ that contains only marginally stable and unstable eigenvalues.

From the system model of P_1 in (2.4), the definition of controller $\mathcal{C}_1$ in (2.5), we may write the system state $\mathbf{x}_1(k)$ as

$$\mathbf{x}_1(k) = F^k\mathbf{x}_1(0) + \sum_{i=0}^{k-1} F^{k-i-1}Gg_1\big(\mathbf{e}_1^i\big). \tag{2.7}$$

With the definition of tracking error $\boldsymbol{\xi}_1(k) = \mathbf{r}(k) - \mathbf{x}_1(k)$, by rearranging the terms in (2.7), we have

$$-F^{-k}\big(\boldsymbol{\xi}_1(k) - \mathbf{r}(k)\big) = \mathbf{x}_1(0) + \sum_{i=0}^{k-1} F^{-i-1}Gg_1\big(\mathbf{e}_1^i\big). \tag{2.8}$$

For bounded reference signals $\mathbf{r}(k)$ and $\mathbb{E}[\boldsymbol{\xi}_1(k)^T\boldsymbol{\xi}_1(k)] < \infty$, from the triangle inequality we have $\mathbb{E}[\mathbf{x}_1(k)^T\mathbf{x}_1(k)] \leq \mathbb{E}[\boldsymbol{\xi}_1(k)^T\boldsymbol{\xi}_1(k)] + \mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$, implying that the system remains stable. From the definition and properties of the mutual information, we can easily show that

$$I(\mathbf{x}_1(0); \mathbf{e}_1^k|\mathbf{r}^k) \geq I(\mathbf{x}_1(0); \mathbf{e}_1^k) = h(\mathbf{x}_1(0)) - h(\mathbf{x}_1(0)|\mathbf{e}_1^k), \tag{2.9}$$

where we have used the fact that $\mathbf{x}_1(0)$ and $\mathbf{r}^k$ are independent. From (2.8) and (P5):

$$\begin{aligned} h(\mathbf{x}_1(0)|\mathbf{e}_1^k) &= h\big(-F^{-k}(\boldsymbol{\xi}_1(k)-\mathbf{r}(k))|\mathbf{e}_1^k\big), \\ &\le h\big(-F^{-k}(\boldsymbol{\xi}_1(k)-\mathbf{r}(k))\big), \qquad (2.10) \\ &\le \frac{1}{2}\log_2\big((2\pi e)^n \det(\mathrm{Cov}\{-F^{-k}(\boldsymbol{\xi}_1-\mathbf{r})\})\big), \qquad (2.11) \\ &= \frac{n}{2}\log_2\big(2\pi e\big)+\frac{1}{2}\log_2\big(\det(F^{-k}(F^{-k})^T)\big)+\frac{1}{2}\log_2\big(\det(\mathrm{Cov}\{\boldsymbol{\xi}_1-\mathbf{r}\})\big), \\ &= \frac{n}{2}\log_2\big(2\pi e\big)-k\sum_i \log_2\left(|\lambda_i(F)|\right)+\frac{1}{2}\log_2\big(\det(\mathrm{Cov}\{\boldsymbol{\xi}_1-\mathbf{r}\})\big), \end{aligned}$$

where (2.10) is due to (P6) and (2.11) is from (P7). Substituting these into (2.9), we obtain

$$\begin{aligned} I(\mathbf{x}_1(0);\mathbf{e}_1^k|\mathbf{r}^k) \ge I(\mathbf{x}_1(0);\mathbf{e}_1^k) &\ge h(\mathbf{x}_1(0))-\frac{n}{2}\log_2(2\pi e) \\ &\quad + k\sum_i \log_2\left(|\lambda_i(F)|\right)-\frac{1}{2}\log_2\left(\det(\mathrm{Cov}\{\boldsymbol{\xi}_1-\mathbf{r}\})\right). \end{aligned}$$

Since $\mathbb{E}[\mathbf{x}_1(k)\mathbf{x}_1(k)^T] < \infty$ and $\mathbf{x}_1 = \mathbf{r}-\boldsymbol{\xi}_1$, we have $\log_2\big(\det(\mathrm{Cov}\{\boldsymbol{\xi}_1-\mathbf{r}\})\big) < \infty$. Finally, if we divide above by k and take the limit $k\to\infty$, then we have $\lim_{k\to\infty}\frac{I(\mathbf{x}_1(0);\mathbf{e}_1^k|\mathbf{r}^k)}{k} \ge \lim_{k\to\infty}\frac{I(\mathbf{x}_1(0);\mathbf{e}_1^k)}{k} \ge \sum_i \log_2(|\lambda_i(F)|)$. If we reintroduce matrix F with some stable eigenvalues, the Lemma follows. □

By applying Lemma 2.4.1 to plant P_2 and with the definition of tracking error $\boldsymbol{\xi}_2$, we have the following corollary.

Corollary 2.4.2. *Consider the closed-loop system in block 2 in Fig. 2.2, where plant P_2 is an LTI system described by (2.4). Assume that the pair (F,G) is controllable and (F,H) is observable. Let L be a transformation matrix such that $\mathbf{x}_2 = L\mathbf{y}_2$. Assume $\mathbb{E}[\hat{\mathbf{y}}_{12}(k)^T\hat{\mathbf{y}}_{12}(k)] < \infty$ for signal $\hat{\mathbf{y}}_{12}(k)$. If $\mathbb{E}[(L\hat{\mathbf{y}}_{12}(k)-\mathbf{x}_2(k))(L\hat{\mathbf{y}}_{12}(k)-\mathbf{x}_2(k))^T] < \infty$, then*

$$\lim_{k\to\infty}\frac{I(\mathbf{x}_2(0);\mathbf{e}_2^k|\hat{\mathbf{y}}_{12}^k)}{k} \ge \lim_{k\to\infty}\frac{I(\mathbf{x}_2(0);\mathbf{e}_2^k)}{k} \ge \sum_i \max\{0,\log_2(|\lambda_i(F)|)\}.$$

Proof. See the proof of Lemma 2.4.1. □

We now examine necessary conditions on the channel rate of CH_{11} to guarantee $\mathbb{E}[\boldsymbol{\xi}_1(k)^T\boldsymbol{\xi}_1(k)] < \infty$ in plant P_1.

Lemma 2.4.3. *Consider the closed-loop system given in block 1 in Fig. 2.2, where the plant P_1 is an LTI system described by (2.4). The channel CH_{11} is a feedback link with rate C_{11}. Assume finite power for reference signal $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$. If $\mathbb{E}[\boldsymbol{\xi}_1(k)^T\boldsymbol{\xi}_1(k)] < \infty$, then*

$$C_{11} \geq I_\infty(\mathbf{r}, \hat{\mathbf{y}}_{11}) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}. \tag{2.12}$$

Proof. By the chain rule (P8) for mutual information , we have

$$I((\mathbf{r}^k, \mathbf{x}_1(0)); \hat{\mathbf{y}}_{11}^k) = I(\mathbf{r}^k, \hat{\mathbf{y}}_{11}^k) + I(\mathbf{x}_1(0); \hat{\mathbf{y}}_{11}^k | \mathbf{r}^k). \tag{2.13}$$

From (P3) and using the fact that $\mathbf{e}_1^k = H\mathbf{r}^k - \hat{\mathbf{y}}_{11}^k$, we have

$$\begin{aligned} I(\mathbf{x}_1(0); \hat{\mathbf{y}}_{11}^k | \mathbf{r}^k) &= h(\hat{\mathbf{y}}_{11}^k | \mathbf{r}^k) - h(\hat{\mathbf{y}}_{11}^k | \mathbf{x}_1(0), \mathbf{r}^k), \\ &= h(\mathbf{e}_1^k | \mathbf{r}^k) - h(\mathbf{e}_1^k | \mathbf{x}_1(0), \mathbf{r}^k), \\ &= I(\mathbf{x}_1(0); \mathbf{e}_1^k | \mathbf{r}^k). \end{aligned} \tag{2.14}$$

Substituting (2.13) into (2.14), we have

$$I(\mathbf{r}^k, \hat{\mathbf{y}}_{11}^k) = I((\mathbf{r}^k, \mathbf{x}_1(0)); \hat{\mathbf{y}}_{11}^k) - I(\mathbf{x}_1(0); \mathbf{e}_1^k | \mathbf{r}^k). \tag{2.15}$$

From Lemma 7.9.2 of [128], for a discrete memoryless channel, we have average data rate $kC_{11} \geq I((\mathbf{r}^k, \mathbf{x}_1(0)); \hat{\mathbf{y}}_{11}^k)$. Hence, from (2.15), we obtain

$$I(\mathbf{r}^k, \hat{\mathbf{y}}_{11}^k) \leq kC_{11} - I(\mathbf{x}_1(0); \mathbf{e}_1^k | \mathbf{r}^k). \tag{2.16}$$

If we divide (2.16) by k and take the limit $k \to \infty$, then the result follows from Lemma 2.4.1. □

From Lemma 2.4.3, we know that for the system P_1 to track $\mathbf{r}(k)$ with finite energy error, the channel data rate of CH_{11} should be at least as large as $\sum_i \max\{0, \log_2(|\lambda_i(F)|)\} + I_\infty(\mathbf{r}, \hat{\mathbf{y}}_{11})$.

The term $I_\infty(\mathbf{r}, \hat{\mathbf{y}}_{11})$ is the average amount of information about $\mathbf{r}$ contained in channel output $\hat{\mathbf{y}}_{11}$ over time. It can be seen from Fig. 2.2 that channel CH_{11} conveys the information of $\mathbf{r}$ and the uncertainty of system P_1 to channel output $\hat{\mathbf{y}}_{11}$. Hence, the channel data rate should be larger than the sum of system uncertainty $\sum_i \max\{0, \log_2(|\lambda_i(F)|)\}$ and the mutual information rate between $\mathbf{r}$ and $\hat{\mathbf{y}}_{11}$. In practice, mutual information rate could be estimated by Monte Carlo methods given a large amount of data [129, 130].

Channels CH_2 and CH_{12}

If we consider $\mathbf{y}_1(k)$ as the reference signal to plant P_2, then block 2 in Fig. 2.2 could be considered a closed-loop system with communication channels in both forward and feedback links. In the following lemma, first we examine necessary conditions on the data rate of feedback channel CH_2 to guarantee that the tracking error has finite second moment for plant P_2 to track $\hat{\mathbf{y}}_{12}(k)$.

Lemma 2.4.4. *Consider the closed-loop system given in block 2 in Fig. 2.2, where the plant P_2 is an LTI system described by (2.4). The channel CH_2 is a feedback link with data rate C_2. Assume that the pair (F, G) is controllable and (F, H) is observable. Let L be a transformation matrix such that $\mathbf{x}_j = L\mathbf{y}_j$ for $j = 1, 2$. Assume $\mathbb{E}[\hat{\mathbf{y}}_{\mathbf{12}}(k)^T \hat{\mathbf{y}}_{12}(k)] < \infty$. If $\mathbb{E}[(L\hat{\mathbf{y}}_{12}(k) - \mathbf{x}_2(k))^T (L\hat{\mathbf{y}}_{12}(k) - \mathbf{x}_2(k))] < \infty$, then*

$$C_2 \geq I_\infty(\hat{\mathbf{y}}_{12}, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}. \tag{2.17}$$

Proof. Similar to the proof of Lemma 2.4.3, and is omitted. □

Lemma 2.4.4 shows that if plant P_2 tracks $\hat{\mathbf{y}}_{12}(k)$ with $\mathbb{E}[(L\hat{\mathbf{y}}_{12}(k) - \mathbf{x}_2(k))^T (L\hat{\mathbf{y}}_{12}(k) - \mathbf{x}_2(k))] < \infty$, then (2.17) should be satisfied. However, our goal is to guarantee $\mathbb{E}[\boldsymbol{\xi}_2(k)^T \boldsymbol{\xi}_2(k)] < \infty$ for plant P_2 to track $\mathbf{y}_1(k)$. Therefore, we need to derive necessary conditions on forward channel CH_{12} and determine the interaction of the two channels in such a system.

In the rest of this section, we first start with Lemma 2.4.5 which shows a lower bound on the rate of forward channel CH_{12} by assuming that feedback channel CH_2 is lossless and has no delays. Later, we will relax this assumption and arrive at our main result on necessary conditions on both channels CH_{12} and CH_2 in Theorem 2.4.7.

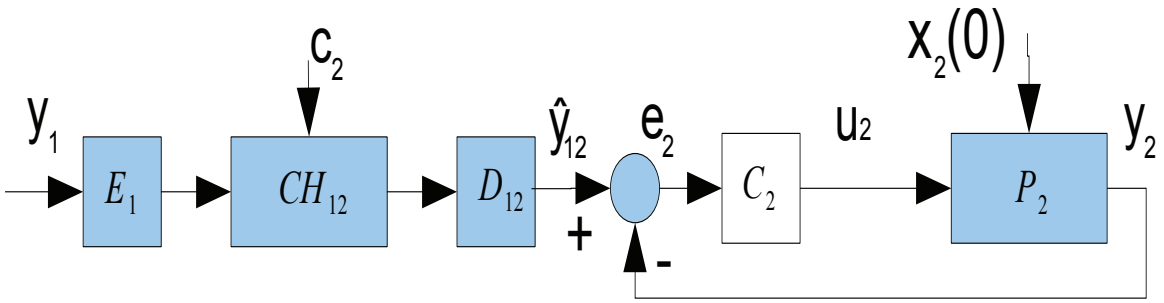


Figure 2.3: Closed-loop system (block 2) with communication channel in forward link and channel CH_2 as lossless with no delays.

Lemma 2.4.5. *Consider the feedback interconnection represented in Fig. 2.3, where the plant P_2 is an LTI system described by (2.4). Assume that the channel CH_2 is lossless and has no delays. Assume that encoder ϵ_1 and decoder D_{12}^k are causal and $\mathbb{E}[\mathbf{y_1}(k)^T\mathbf{y}_1(k)] < \infty$. Assume that the pair (F, G) is controllable and (F, H) is observable. Let L be a transformation matrix such that $\mathbf{x}_j = L\mathbf{y}_j$ for $j = 1, 2$. If $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$, then*

$$C_{12} \geq I_\infty(\mathbf{y}_1, \mathbf{y}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\} - h(\mathbf{x}_2(0)), \tag{2.18}$$

where C_{12} represents the rate of channel CH_{12}. In addition, channel noise $\mathbf{c}_{12} = 0$, then (2.18) is given by

$$C_{12} \geq I_\infty(\hat{\mathbf{y}}_{12}, \mathbf{y}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\} - h(\mathbf{x}_2(0)), \tag{2.19}$$

Proof. From the chain rule (P8) for mutual information,

$$I((\mathbf{x}_2(0), \mathbf{y}_1^k); \mathbf{y}_2^k) = I(\mathbf{y}_1^k; \mathbf{y}_2^k) + I(\mathbf{x}_2(0); \mathbf{y}_2^k | \mathbf{y}_1^k). \tag{2.20}$$

From the system model (2.4) of P_2 and the definition (2.5) of controller $\mathcal{C}_2$, we may write the output $\mathbf{y}_2(k)$ as

$$\begin{aligned}\mathbf{y}_2(k) &= HF^k\mathbf{x}_2(0) + H\sum_{i=0}^{k-1} F^{k-i-1}Gg_2\big(\hat{\mathbf{y}}_{12}^i - \mathbf{y}_2^i\big),\\ &= \hat{g}_2(\mathbf{x}_2(0), \hat{\mathbf{y}}_{12}(k)).\end{aligned} \tag{2.21}$$

This shows that $\mathbf{y}_2(k)$ is a function of initial state $\mathbf{x}_2(0)$ and the reference signal $\hat{\mathbf{y}}_{12}$. From (P3) and (2.21), we obtain

$$\begin{aligned}I\big((\mathbf{x}_2(0), \mathbf{y}_1^k); \mathbf{y}_2^k\big) \leq & I\big((\mathbf{x}_2(0), \mathbf{y}_1^k); \mathbf{q}^k\big),\\ =& I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k) + I\big(\mathbf{x}_2(0), (\mathbf{x}_2(0), \cdots, \mathbf{x}_2(0))\big),\\ =& I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k) + kh(\mathbf{x}_2(0)),\end{aligned} \tag{2.22}$$

where $\mathbf{q}^k = \{\mathbf{q}(1), \cdots, \mathbf{q}(k)\}$, $\mathbf{q}(k) = (\mathbf{x}_2(0), \hat{\mathbf{y}}_{12}(k))$ and the second step results from the independence between $\mathbf{x}_2(0)$ and $(\mathbf{y}_1^k, \hat{\mathbf{y}}_{12}^k)$. From (P8), we have

$$\begin{aligned}I((\mathbf{y}_1^k, \hat{\mathbf{y}}_{12}^k); \mathbf{y}_2^k) &= I(\mathbf{y}_1^k; \mathbf{y}_2^k) + I(\hat{\mathbf{y}}_{12}^k; \mathbf{y}_2^k|\mathbf{y}_1^k),\\ &= I(\hat{\mathbf{y}}_{12}^k; \mathbf{y}_2^k) + I(\mathbf{y}_1^k; \mathbf{y}_2^k|\hat{\mathbf{y}}_{12}^k).\end{aligned}$$

Since $I(\hat{\mathbf{y}}_{12}^k; \mathbf{y}_2^k|\mathbf{y}_1^k) = 0$ due to (P4) and $\mathbf{c}_{12} = 0$ and $I(\mathbf{y}_1^k; \mathbf{y}_2^k|\hat{\mathbf{y}}_{12}^k) \geq 0$ due to (P1), we have

$$I(\mathbf{y}_1^k; \mathbf{y}_2^k) \geq I(\hat{\mathbf{y}}_{12}^k; \mathbf{y}_2^k). \tag{2.23}$$

From the definition of mutual information, we further have

$$\begin{aligned}I(\mathbf{x}_2(0); \mathbf{y}_2^k|\mathbf{y}_1^k) &= h(\mathbf{x}_2(0)|\mathbf{y}_1^k) - h(\mathbf{x}_2(0)|\mathbf{y}_2^k, \mathbf{y}_1^k),\\ &= h(\mathbf{x}_2(0)|\hat{\mathbf{y}}_{12}^k) - h(\mathbf{x}_2(0)|\mathbf{y}_2^k, \hat{\mathbf{y}}_{12}^k) &(2.24)\\ &= I(\mathbf{x}_2(0); \mathbf{y}_2^k|\hat{\mathbf{y}}_{12}^k),\\ &= h(\mathbf{y}_2^k|\hat{\mathbf{y}}_{12}^k) - h(\mathbf{y}_2^k|\mathbf{x}_2(0), \hat{\mathbf{y}}_{12}^k),\\ &= h(\mathbf{e}_2^k|\hat{\mathbf{y}}_{12}^k) - h(\mathbf{e}_2^k|\mathbf{x}_2(0), \hat{\mathbf{y}}_{12}^k), &(2.25)\\ &= I(\mathbf{x}_2(0); \mathbf{e}_2^k|\hat{\mathbf{y}}_{12}^k), &(2.26)\end{aligned}$$

where (2.24) is due to the independence between $\mathbf{x}_2(0)$ and $(\mathbf{y}_1^k, \hat{\mathbf{y}}_{12}^k)$ and (2.25) is due to (P5) and the fact that $\mathbf{e}_2^k = \hat{\mathbf{y}}_{12}^k - \mathbf{y}_2^k$. Substitution of (2.22), (2.23) and (2.26) into (2.20) results in the following:

$$I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k) + kh(\mathbf{x}_2(0)) \geq I(\hat{\mathbf{y}}_{12}^k; \mathbf{y}_2^k) + I(\mathbf{x}_2(0); \mathbf{e}_2^k | \hat{\mathbf{y}}_{12}^k).$$

We have average channel data rate $kC_{12} \geq I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k)$. By dividing above by k and taking the limit $k \to \infty$, the result follows by using Corollary 2.4.2. □

Remark 2.4.6. *From Lemma 2.4.5, it could be seen that if $C_{12} < \sum_i \max\{0, \log_2(|\lambda_i(F)|)\} - h(\mathbf{x}_2(0))$, the channel can not convey information at a high enough rate to match the speed of the system dynamics such that the reference signal does not provide any information related to the feedback signal, rendering feedback useless. By comparing Lemma 2.4.4 and Lemma 2.4.5, we know that there is one more term $-h(\mathbf{x}_2(0))$ in (2.19), which is due to the fact that $\mathbf{x}_2(0)$ passes through channel CH$_2$ but does not go through channel CH$_{12}$. When calculating the channel data rate, the information of $\mathbf{x}_2(0)$ is taken into account in C_2 but not in C_{12}. If we assume $\mathbf{x}_2(0)$ is not a random variable but a deterministic one, there is no uncertainty in $\mathbf{x}_2(0)$ such that $h(\mathbf{x}_2(0)) = 0$ and the bounds on C_2 and C_{12} are the same.*

By combining Lemma 2.4.4 and Lemma 2.4.5, we provide the following theorem which states the necessary conditions on channels CH$_{12}$ and CH$_2$ together to guarantee $\mathbb{E}[\boldsymbol{\xi}_2(k)^T \boldsymbol{\xi}_2(k)] < \infty$ for tracking of $\mathbf{y}_1(k)$ by plant P_2.

Theorem 2.4.7. *Consider block 2 represented in Fig. 2.2, where plant P_2 is an LTI system described by (2.4). Assume that encoders ϵ_1, ϵ_2 and decoders D_{12}^k, D_2^k are causal and $\mathbb{E}[\mathbf{r}(k)^T \mathbf{r}(k)] < \infty$. Assume that the pair (F, G) is controllable and (F, H) is observable. Let L be a transformation*

matrix such that $\mathbf{x}_j = L\mathbf{y}_j$ *for* $j = 1, 2$. *If* $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$, *then* [3]

$$C_2 \geq I_\infty(\hat{\mathbf{y}}_{12}, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}, \text{ and}$$
$$C_{12} \geq I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\} - h(\mathbf{x}_2(0)), \tag{2.27}$$

where C_2 *and* C_{12} *are the channel rates of* CH_2, CH_{12}.

Proof. The first equation in (2.27) results directly from Lemma 2.4.4. In the following, we provide the proof for the second equation. From the chain rule (P8) for mutual information, we obtain

$$I((\mathbf{x}_2(0), \mathbf{y}_1^k); \hat{\mathbf{y}}_2^k) = I(\mathbf{y}_1^k; \hat{\mathbf{y}}_2^k) + I(\mathbf{x}_2(0); \hat{\mathbf{y}}_2^k | \mathbf{y}_1^k). \tag{2.28}$$

From (2.26) and using the fact that $\mathbf{e}_2^k = \hat{\mathbf{y}}_{12}^k - \hat{\mathbf{y}}_2^k$, we have

$$I(\mathbf{x}_2(0); \hat{\mathbf{y}}_2^k | \mathbf{y}_1^k) = I(\mathbf{x}_2(0); \mathbf{e}_2^k | \hat{\mathbf{y}}_{12}^k). \tag{2.29}$$

Using the properties of mutual information, we have

$$I((\mathbf{x}_2(0), \mathbf{y}_1^k); \hat{\mathbf{y}}_2^k) \leq I((\mathbf{x}_2(0), \mathbf{y}_1^k); (\mathbf{y}_2^k, \mathbf{c}_2^k)), \tag{2.30}$$
$$= I((\mathbf{x}_2(0), \mathbf{y}_1^k); \mathbf{y}_2^k), \tag{2.31}$$
$$\leq I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k) + kh(\mathbf{x}_2(0)), \tag{2.32}$$

where (2.30) results from (P3) and $\hat{\mathbf{y}}_2^k$ is a function of $\mathbf{y}_2^k$ and $\mathbf{c}_2^k$, (2.30) is due to the independence between $\mathbf{c}_2^k$ and $(\mathbf{x}_2(0), \mathbf{y}_1^k)$ and (2.32) results from (2.22). Substitution of (2.29) and (2.32) into (2.28) results in the following

$$I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k) + kh(\mathbf{x}_2(0)) \geq I(\mathbf{y}_1^k; \hat{\mathbf{y}}_2^k) + I(\mathbf{x}_2(0); \mathbf{e}_2^k | \hat{\mathbf{y}}_{12}^k).$$

We have the average channel rate as $kC_{12} \geq I(\mathbf{y}_1^k; \hat{\mathbf{y}}_{12}^k)$. Hence, by dividing above by k and taking the limit $k \to \infty$, the result follows by using Corollary 2.4.2. □

[3] The calculation of information rate in practice is difficult due to enormous computation. However, it still could be closely estimated by Monte Carlo techniques given a large enough amount of data [129, 130].

In Theorem 2.4.7, it can be seen that when two channels appear as forward link (CH_{12}) and feedback link (CH_2) as in block 2 in Fig. 2.2, the rate of the forward channel is affected by the existence of the feedback channel, if the feedback channel is noisy where we have $I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2)$ in (2.27) instead of $I_\infty(\mathbf{y}_1, \mathbf{y}_2)$ in (2.18). The physical meaning of Theorem 2.4.7 is that the channels CH_{12} and CH_2 should convert information at a high enough rate not just to guarantee system stability by stabilizing the unstable poles of system matrix ($\sum_i \max\{0, \log_2(|\lambda_i(F)|)\}$), but also ensure effective tracking by providing related information between reference signal and feedback signal, which is represented by the mutual information $I_\infty(\hat{\mathbf{y}}_{12}, \hat{\mathbf{y}}_2)$ and $I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2)$.

2.4.2 Necessary conditions on cascade channel made of CH_{12} and CH_2

Theorem 2.4.7 shows necessary conditions on data rate of each individual link for tracking in plant P_2 with finite energy tracking error. It also shows the interaction between forward and feedback channels in this networked feedback system. However, a general overview and abstraction of the necessary conditions on both channels for tracking in such a system is still needed. With regards to information flow, the forward and feedback channels could be connected in a cascade manner. By considering forward and feedback channels together as one cascade channel, we could provide a lower bound on the rate of the cascade channel for tracking in plant P_2 such that $\mathbb{E}[\boldsymbol{\xi}_2(k)^T \boldsymbol{\xi}_2(k)] < \infty$.

We may reformulate the structure of block 2 in Fig. 2.2 as shown in Fig. 2.4 [29]. In Fig. 2.4, two channels are connected in cascade with feedback from the output of the second channel to the intermediate node. The first channel $\text{CH}_{12,\text{new}}$ consists of CH_{12} and the lossless link that transmits $\mathbf{x}_2(0)$. The encoder input of the first channel is denoted as $(\mathbf{y}_1, \mathbf{x}_2(0))$ and the decoder output is denoted by $(\hat{\mathbf{y}}_{12}, \mathbf{x}_2(0))$, since $\mathbf{x}_2(0)$ is not affected by the channel noise. For the second channel CH_2, we denote the encoder input as $(\hat{\mathbf{y}}_{12}, \mathbf{x}_2(0))$ and decoder output as $\hat{\mathbf{y}}_2$. Here we consider controller $\mathcal{C}_2$, plant P_2 and encoder ϵ_2 as a macro encoder for the second channel.

In this formulation, we reconsider the two channels $\text{CH}_{12,\text{new}}$, CH_2 described in Fig. 2.4 as one

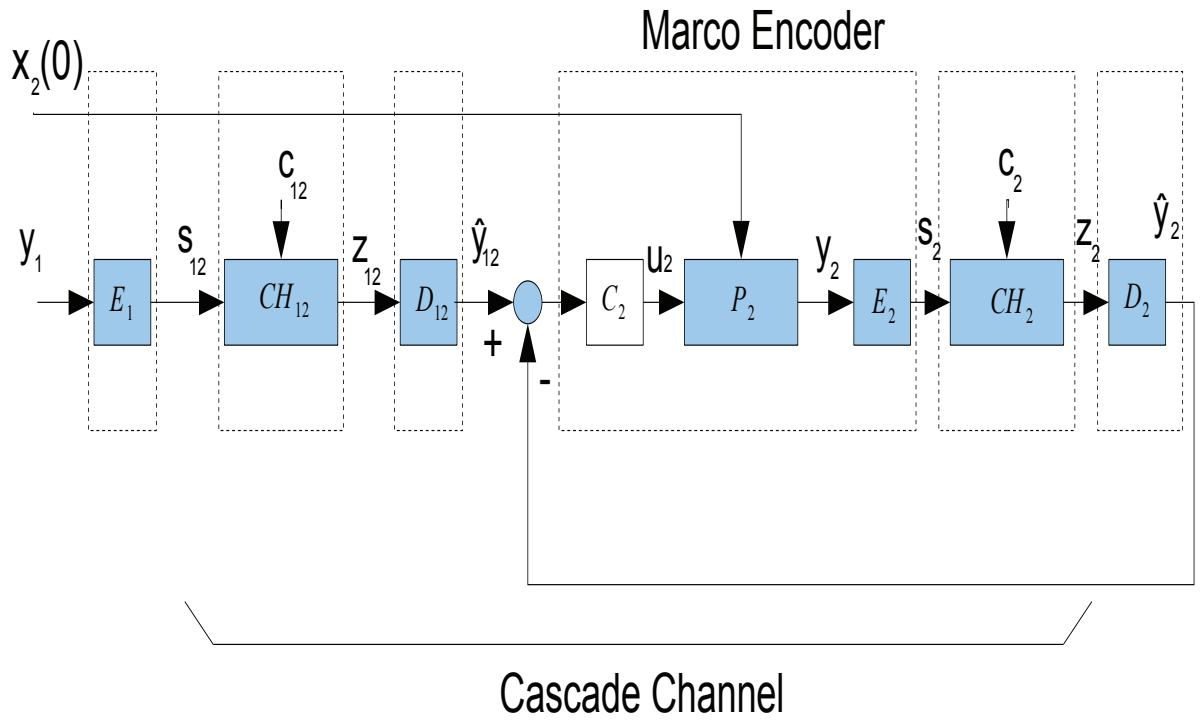


Figure 2.4: Closed-loop system (block 2) with communication channels in forward and feedback links.

cascade channel CH_{cas} with encoder ϵ_1, decoder D_2 and the components in between as the new channel. Here we want to find the minimum channel rate for the cascade channel CH_{cas} for plant P_2 to track $\mathbf{y}_1(k)$ with $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$.

In [29], a similar reformation of closed-loop system with feedback channel is considered. Here we extend the results in [29] to a cascade channel in which both forward and feedback channels are considered. In order to proceed with our results, we first modify Theorem 2.1 in [29] and state it here as Lemma 2.4.8.

Lemma 2.4.8. *Consider the closed-loop system given in Fig. 2.4, where the plant P_2 is an LTI system described by (2.4). Assume that the encoders and decoders for the two channels CH_{12} and CH_2 are causal operators. Assume that the pair (F, G) is controllable and (F, H) is observable.*

Let L be a transformation matrix such that $\mathbf{x}_j = L\mathbf{y}_j$ for $j = 1, 2$. Let the following assumptions hold:

(A1) The decoder D_2 for the second channel CH$_2$ satisfies: $\forall k > a, \hat{\mathbf{y}}_2^{a+1,k} = D_2^k(\hat{\mathbf{y}}_2^{1,a}, \mathbf{z}_2^k)$ for some $a \in \mathcal{N}_+$ and a sequence of functions D_2^k, where $\hat{\mathbf{y}}_2^{a+1,k} = \{\hat{\mathbf{y}}_2(a+1), \cdots, \hat{\mathbf{y}}_2(k)\}$ and the output of the decoder D_2 is based on all the received values from channel $\mathbf{z}_2^k$ and the previous output of decoder $\hat{\mathbf{y}}_2^{1,a}$.

(A2) The fading memory condition $\limsup_{k\to\infty} \frac{1}{k} I(\hat{\mathbf{y}}_2^{1,a}; \mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k) = 0$ holds.

Under the above conditions, the following is true:

$$\limsup_{k\to\infty} \frac{1}{k} I(\mathbf{x}_2(0), \mathbf{y}_1^k|\mathbf{z}_2^k) \leq I_\infty(\mathbf{s}_{12} \to \mathbf{z}). \tag{2.33}$$

Proof. We separate the proof into two parts.

1) First, using (P2) and (P10) we can write the following equality, for any given $i \in \{1, \cdots, k\}$:

$$I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}) \leq I(\mathbf{z}_2(i); \mathbf{s}_{12}^i|\mathbf{z}_2^{i-1}) + I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i). \tag{2.34}$$

Now notice that (P2) allows us to rewrite:

$$I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i) = I((\mathbf{z}_2^i, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})) - I((\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})). \tag{2.35}$$

But, from (P3), we know that

$$I((\mathbf{z}_2^i, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})) = I((\mathbf{n}(i), \mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})).$$

where $\mathbf{n}(i)$ represents the additive components in the cascade channel CH$_{\text{cas}}$, including additive noises $\mathbf{c}_{12}(i)$, $\mathbf{c}_2(i)$ and channel output feedback $\hat{\mathbf{y}}_2(i)$. Then, by chain rule, we have

$$\begin{aligned} I((\mathbf{z}_2^i, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})) &= I((\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})) \\ &\quad + I(\mathbf{n}(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|(\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i)). \end{aligned} \tag{2.36}$$

Since $\mathbf{n}(i)$ is independent of $(\mathbf{x}_2(0), \mathbf{y}_1^{i-1})$ given $(\mathbf{s}_{12}^i, \mathbf{z}_2^{i-1})$, we have $I(\mathbf{n}(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|(\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i)) = 0$. Then,

$$I((\mathbf{z}_2^i, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})) = I((\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})). \tag{2.37}$$

By making use of (2.35) and (2.37) we infer that $I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}, \mathbf{s}_{12}^i) = 0$. Together with (P1) and (2.34), this leads to:

$$I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}) \leq I(\mathbf{z}_2(i); \mathbf{s}_{12}^i|\mathbf{z}_2^{i-1}). \tag{2.38}$$

From causality (A1), $\mathbf{y}_1^{i,k}$ is independent of $(\mathbf{x}_2(0), \mathbf{y}_1^{i-1}, \mathbf{z}_2^i)$ implying

$$I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{k})|\mathbf{z}_2^{i-1}) = I(\mathbf{z}_2(i); (\mathbf{x}_2(0), \mathbf{y}_1^{i-1})|\mathbf{z}_2^{i-1}). \tag{2.39}$$

Substituting (2.39) in (2.38) and summing over i from $i = 1$ to $i = k$, then we have

$$I(\mathbf{z}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)) \leq I(\mathbf{s}_{12}^k \to \mathbf{z}_2^k). \tag{2.40}$$

2) Second, by using (P2) and (P10), we have the following inequality

$$I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)) \leq I(\mathbf{z}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)) + I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k). \tag{2.41}$$

From (P2) and assumption (A1), we obtain the following:

$$I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k) = I(\hat{\mathbf{y}}_2^{a+1,k}; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k, \hat{\mathbf{y}}_2^a) + I(\hat{\mathbf{y}}_2^a; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k).$$

From (P4), we have $I(\hat{\mathbf{y}}_2^{a+1,k}; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k, \hat{\mathbf{y}}_2^a) = 0$. Then,

$$I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k) = I(\hat{\mathbf{y}}_2^a; (\mathbf{x}_2(0), \mathbf{y}_1^k)|\mathbf{z}_2^k). \tag{2.42}$$

By substitution of (2.42) in (2.41) and using the assumption (A2), we obtain:

$$\limsup_{k\to\infty} \frac{1}{k} I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)) \leq \limsup_{k\to\infty} \frac{1}{k} I(\mathbf{z}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)),$$

which, together with (2.40), completes the proof. □

By the definitions of channel rate and directed mutual information, from Lemma 2.4.8, we have $C_{\text{cas}} \geq I_\infty(\mathbf{s}_{12} \to \mathbf{z}_2) \geq \limsup_{k\to\infty} \frac{1}{k} I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k))$. Then, we have the following:

$$C_{\text{cas}} \geq \limsup_{k\to\infty} \frac{1}{k} I(\hat{\mathbf{y}}_2^k; (\mathbf{x}_2(0), \mathbf{y}_1^k)). \tag{2.43}$$

Since $I((\mathbf{y}_1^k, \mathbf{x}_2(0)); \hat{\mathbf{y}}_2^k) = I(\mathbf{y}_1^k, \hat{\mathbf{y}}_2^k) + I(\mathbf{x}_2(0); \hat{\mathbf{y}}_2^k | \mathbf{y}_1^k)$, from (P3) and the fact that $\mathbf{e}_2^k = \hat{\mathbf{y}}_{12}^k - \hat{\mathbf{y}}_2^k = \hat{f}(\mathbf{y}_1^k) - \hat{\mathbf{y}}_2^k$, we have

$$\begin{aligned} I(\mathbf{x}_2(0); \hat{\mathbf{y}}_2^k | \mathbf{y}_1^k) &= h(\hat{\mathbf{y}}_2^k | \mathbf{y}_1^k) - h(\hat{\mathbf{y}}_2^k | \mathbf{x}_2(0), \mathbf{y}_1^k) \\ &= h(\mathbf{e}_2^k | \mathbf{y}_1^k) - h(\mathbf{e}_2^k | \mathbf{x}_2(0), \mathbf{y}_1^k); \\ &= I(\mathbf{x}_2(0); \mathbf{e}_2^k | \mathbf{y}_1^k). \end{aligned} \tag{2.44}$$

From (2.44), we have

$$I(\mathbf{y}_1^k, \hat{\mathbf{y}}_2^k) = I((\mathbf{y}_1^k, \mathbf{x}_2(0)); \hat{\mathbf{y}}_2^k) - I(\mathbf{x}_2(0); \mathbf{e}_2^k | \mathbf{y}_1^k). \tag{2.45}$$

By Corollary 2.4.2 and dividing (2.45) by k and taking the limit $k \to \infty$, we obtain

$$I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) \leq \limsup_{k\to\infty} \frac{1}{k} I((\mathbf{y}_1^k, \mathbf{x}_2(0)); \hat{\mathbf{y}}_2^k) - \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}. \tag{2.46}$$

Substituting (2.43) in (2.46), we have

$$I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) \leq C_{\text{cas}} - \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}.$$

This result may be summarized in the following theorem:

Theorem 2.4.9. *Consider the system formulation given in Fig. 2.4, where the plant P_2 is an LTI system described by (2.4). Consider the cascade channel, which is the combination of channels CH_{12} and CH_2. Assume that the pair (F, G) is controllable and (F, H) is observable. Let L be a transformation matrix such that $\mathbf{x}_j = L\mathbf{y}_j$ for $j = 1, 2$. Assume that $\mathbb{E}[\mathbf{r}(k)^T \mathbf{r}(k)] < \infty$. If $\mathbb{E}[\boldsymbol{\xi}_2(k)^T \boldsymbol{\xi}_2(k)] < \infty$, then*

$$C_{cas} \geq I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F)|)\}. \tag{2.47}$$

Remark 2.4.10. *Theorem 2.4.9 shows that the rate of the cascade channel which includes CH_{12} and CH_2 is lower bounded by the mutual information between $\mathbf{y}_1^k$ and $\hat{\mathbf{y}}_2^k$ and a function of the unstable poles of plant P_2. From the data processing inequality, the rate of cascade channel is less than the rate of each component channel [131]. Since channels $CH_{12,new}$ and CH_2 are in cascade connection as in Fig. 2.4, we have that $C_{12,new} \geq C_{cas}$ and $C_2 \geq C_{cas}$, where $C_{12,new}$ is the rate of channel $CH_{12,new}$. Since the link that transmits $\mathbf{x}_2(0)$ is lossless, the channel rate $C_{12,new} = C_{12} + h(\mathbf{x}_2(0))$. The result in Theorem 2.4.9 is also confirmed by Theorem 2.4.7.*

Theorem 2.4.9 provides a guideline for communication channel design in networked feedback systems by giving a lower bound on the required overall channel rate. In practice, we could adjust the rate of each component channel and the channel orderings to optimize the overall channel rate. Theorem 2.4.9 shows that the rate of the cascade channel $I_\infty((\mathbf{y}_1, \mathbf{x}_2(0)); \hat{\mathbf{y}}_2)$ can be estimated in two parts. The first part $I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2)$ could be directly measured by partial input and output relation. The second term $I_\infty(\mathbf{x}_2(0), \hat{\mathbf{y}}_2|\mathbf{y}_1)$ which is caused by internal information loss in the channel could be estimated by the system parameters $\sum_i \max\{0, \log_2(|\lambda_i(F)|)\}$.

2.5 Extensions to Different System Models

In this section, we consider the extension of above results to tracking in a leader-follower system in which the leader system and follower system have different system models as follows:

$$\begin{aligned}
\mathbf{x}_1(k+1) &= F_1\mathbf{x}_1(k) + G_1\hat{\mathbf{u}}_1(k),\\
\mathbf{y}_1(k) &= H_1\mathbf{x}_1(k),\\
&\text{and}\\
\mathbf{x}_2(k+1) &= F_2\mathbf{x}_2(k) + G_2\hat{\mathbf{u}}_2(k),\\
\mathbf{y}_2(k) &= H_2\mathbf{x}_2(k), \qquad k \geq 0,
\end{aligned} \tag{2.48}$$

where the states $\mathbf{x}_1(k)$ and $\mathbf{x}_2(k)$ take values in $\mathbb{R}^n$ and the received control inputs $\hat{\mathbf{u}}_1(k)$ and $\hat{\mathbf{u}}_2(k)$ take values in $\mathbb{R}^r$. The initial states $\mathbf{x}_1(0)$ and $\mathbf{x}_2(0)$ are zero mean Gaussian random variables with

covariance matrices Σ_{01} and Σ_{02}, respectively. The states are observed by sensors that generate the measurements $\mathbf{y}_1(k)$ and $\mathbf{y}_2(k)$ taking values in $\mathbb{R}^q$.

In order to comply to the same formulation as in Section 2.3, the terms $\mathbf{e}_1$ and $\mathbf{e}_2$ are defined as $\mathbf{e}_1 = L_1\mathbf{r} - \hat{\mathbf{y}}_1$ and $\mathbf{e}_2 = L_2^{-1}L_1\hat{\mathbf{y}}_{12} - \hat{\mathbf{y}}_2$, where L_1 and L_2 are invertible transformation matrices and $\mathbf{x}_1 = L_1\mathbf{y}_1$ and $\mathbf{x}_2 = L_2\mathbf{y}_2$. By following the derivation of above lemmas and theorems, similar results could be obtained for tracking in leader-follower system where the leader system and follower system have different system models. Due to space limitation, we only list the main results here and the proofs follow the same derivation as for Theorems 2.4.7 and 2.4.9.

Theorem 2.5.1. *Consider block 2 represented in Fig. 2.2, where plant P_2 is an LTI system described by (2.48). Assume that encoders ϵ_1, ϵ_2 and decoders D_{12}^k, D_2^k are causal and that $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$. Assume that the pair (F_j, G_j) is controllable and (F_j, H_j) is observable. Let L_j be an invertible transformation matrix such that $\mathbf{x}_j = L_j\mathbf{y}_j$ for $j = 1, 2$. If $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$, then*

$$\begin{aligned} C_2 &\geq I_\infty(\hat{\mathbf{y}}_{12}, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F_2)|)\}, \text{ and} \\ C_{12} &\geq I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F_2)|)\} - h(\mathbf{x}_2(0)), \end{aligned} \tag{2.49}$$

where C_2 and C_{12} are the channel rates of CH_2, CH_{12}.

Theorem 2.5.2. *Consider the system formulation given in Fig. 2.4, where the plant P_2 is an LTI system described by (2.48). Consider the cascade channel, which is the combination of channels CH_{12} and CH_2. Assume that the pair (F_j, G_j) is controllable and (F_j, H_j) is observable. Let L_j be an invertible transformation matrix such that $\mathbf{x}_j = L_j\mathbf{y}_j$ for $j = 1, 2$. Assume that $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] < \infty$. If $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$, then*

$$C_{cas} \geq I_\infty(\mathbf{y}_1, \hat{\mathbf{y}}_2) + \sum_i \max\{0, \log_2(|\lambda_i(F_2)|)\}. \tag{2.50}$$

2.6 Examples and Simulations

Leader-follower system is defined as a dynamic system in which multiple agents are connected in such a way that followers are controlled or influenced by the behaviors of leaders. In such system, each follower will keep track of the state or output of the leader and generate its own output based on the received information. In this section, we show simulations of a leader-follower system and demonstrate the necessity of above derived conditions for tracking. We consider the close-loop system as in block 2 in Fig. 2.2. Channels CH_{12} and CH_2 are assumed to be erasure channels with limited data transmission rates R_{12} and R_2 and packet loss erasure probabilities of p_{12} and p_2, respectively. The average data rate of erasure channel is given by $C_i = R_i(1 - p_i)$ [128]. The reference signal satisfies $\mathbb{E}[\mathbf{r}(k)^T\mathbf{r}(k)] \leq 10^3$.

We consider a two-part encoder-decoder scheme as follows [57]: encoder ϵ_i converts the input to its binary form, truncates the binary representation to its R_i most significant bits, encapsulates the bits in a packet and sends the packet through the channel. If the packet is received, the decoder D_i extracts the bits in the packet and converts them to its real number representation. Otherwise, the decoder will assume that a zero was sent and outputs zero. The scheme also assumes that the decoder knows exactly the operation of the encoder and that both have access to control signal. Consider system equation of plant P_2 with a simple control law as $\mathbf{x}_2(k+1) = 16\mathbf{x}_2(k) + \mathbf{u}_2(k)$, $\mathbf{y}_2(k) = 15\mathbf{x}_2(k)$ and $\mathbf{u}_2(k) = \hat{\mathbf{y}}_{12}(k) - 1.07\hat{\mathbf{y}}_2(k)$. The control law will drive system state to $\mathbf{x}_2(k) = \mathbf{r}(k)$ if the two channels are lossless and have no delays. The initial state $\mathbf{x}_2(0)$ is Bernoulli distributed with success probability $p_{x_2} = 0.5$.

Our necessary conditions were given in terms of the mutual information rate, which is difficult to compute directly. However, our results impose limits to guarantee that $\mathbb{E}[\boldsymbol{\xi}_2(k)^T\boldsymbol{\xi}_2(k)] < \infty$. Since $\mathbb{E}[\boldsymbol{\xi}_2^T\boldsymbol{\xi}_2] = \frac{1}{2\pi}\int_0^\pi \Phi_2(\omega)d\omega$, where $\Phi_2(\omega)$ is the power spectral density of $\boldsymbol{\xi}_2(k)$ [132], we may plot the power spectral density of $\boldsymbol{\xi}_2$ to estimate $\mathbb{E}[\boldsymbol{\xi}_2^T\boldsymbol{\xi}_2]$. From Theorem 2.4.7, we know that lower bounds on channel data rate C_2 and C_{12} are 4 bits/timestep, for the above assumed system. Figures 2.5 and 2.6 show the power spectral density $\Phi_2(\omega)$ of the tracking error $\boldsymbol{\xi}_2(k)$ of

the follower system when only one channel satisfies these necessary conditions. It can be seen that the power spectral density $\Phi_2(\omega)$ is unbounded at every $\omega \in [0, \pi]$. Then, the average power spectrum over an area of $[0, \pi]$ is unbounded. From the above equation, we know that $\mathbb{E}[\boldsymbol{\xi}_2^T \boldsymbol{\xi}_2]$ is no longer finite. However, if the lower bounds are satisfied by both channels as assumed in Fig. 2.7, the power spectral density is finite. Hence, $\mathbb{E}[\boldsymbol{\xi}_2^T \boldsymbol{\xi}_2]$ stays bounded.

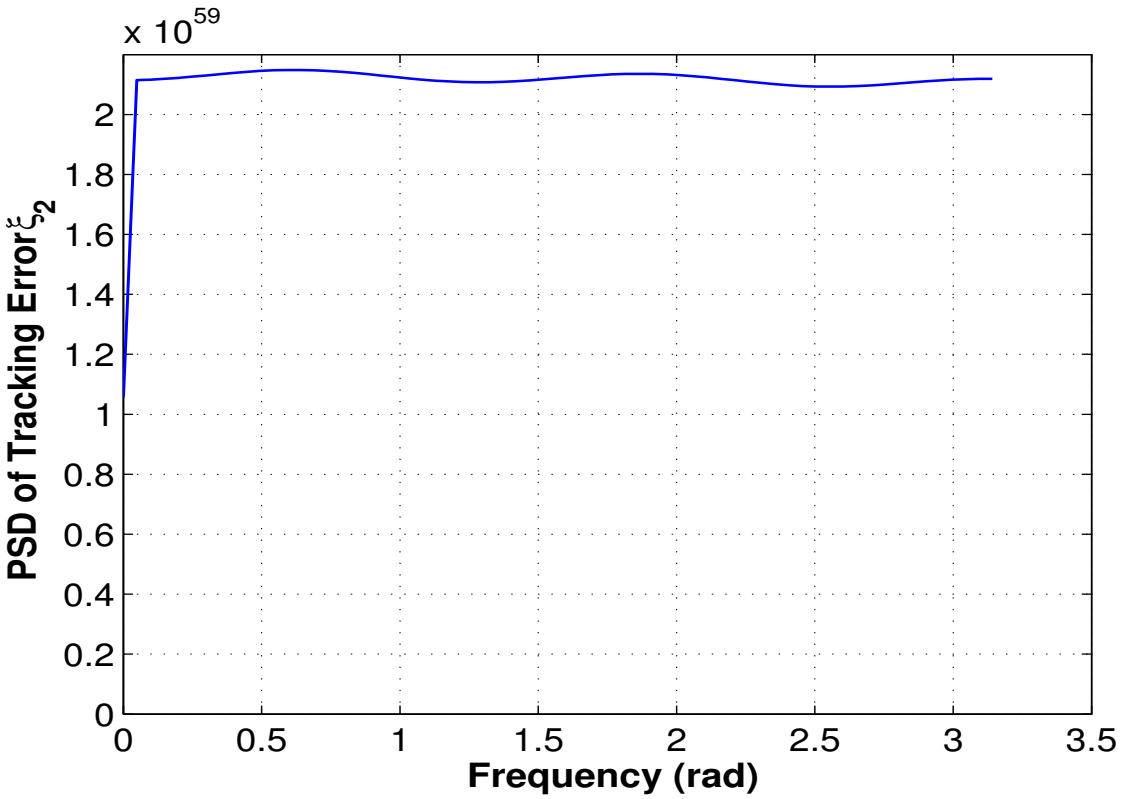


Figure 2.5: Example with erasure channels: $C_{12} = 2.5$, $C_2 = 4$.

2.7 Conclusions

In this chapter, we considered tracking in leader-follower systems under communication constraints, where the system components are distributed and connected over communication links with finite data rates. We provided lower bounds on the channel rate of each communication link

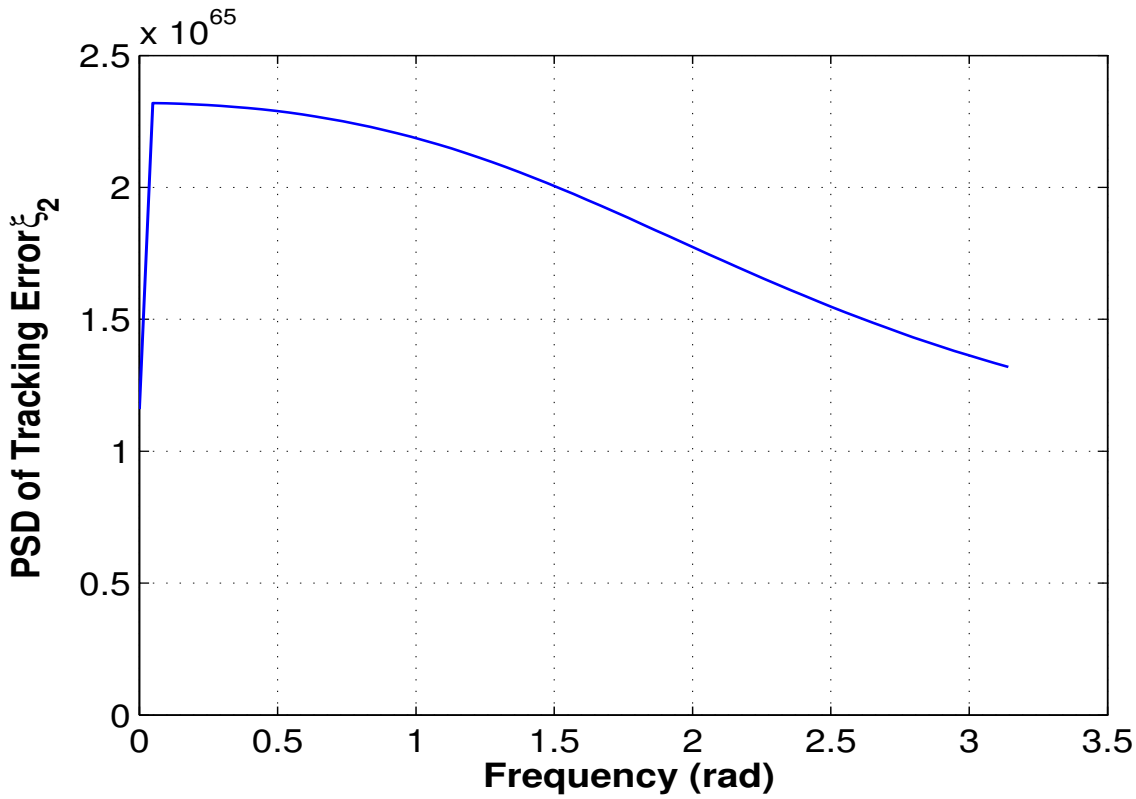


Figure 2.6: Example with erasure channels: $C_{12} = 4$, $C_2 = 2.3$.

as necessary conditions for tracking in such a leader-follower system. We also showed examples to demonstrate our results. The results in this work provide fundamental limitations in terms of information quantities on communication links which can have important roles on control design in leader-follower systems. Limitations in both overall channel and individual channel are provided and it should be taken into account for designing new control system with communication constraints. Our future work is to extend the leader-follower system to more general framework in which multiple leaders and followers are interconnected as a network with more general graph topologies.

In the next chapter, we consider distributed demand-side management for smart grid by optimizing the aggregate load of users with DG through user energy usage coordination. Under the cooperative user and inter-user communication assumption, we propose a distributed control

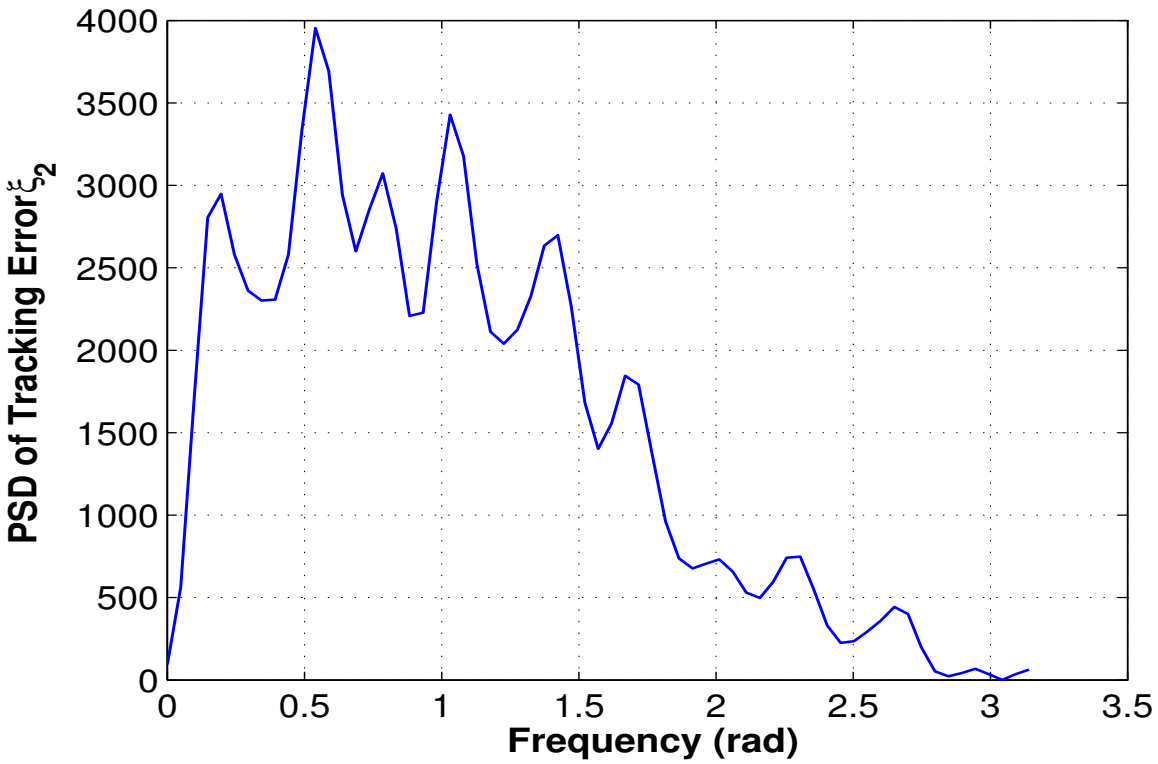


Figure 2.7: Example with erasure channels: $C_{12} = 4$, $C_2 = 9$.

based energy scheduling method to improve the demand shape profile and reduce energy loss. We will consider possible delay in communications, sensing and actuation and propose a distributed networked control strategy with prediction algorithm to solve the problem.

Chapter 3

Distributed Demand-side Management for Mircogrid through Cooperative User Coordination

3.1 Introduction

In this chapter, we consider distributed demand-side management for microgrids by optimizing the aggregate load of users with RDG through user energy usage coordination. Under the assumption of user cooperation and inter-user communications, we propose a distributed networked control strategy to improve the demand shape profile and maintain supply-demand balance. The proposed approach is based on distributed control in which users exchange their own information with other users via inter-user communications and apply optimal control to adjust their energy usage. By considering possible delay in communications between users and management system, we propose a distributed networked control strategy with prediction to solve the aggregate user load optimization problem with same and different communication delays.

The rest of this chapter is organized as follows: Section 3.2 proposes a scalable, hierarchical

decision-making and control architecture for the smart grid. The mirogrid level control framework of the proposed architecture further discussed in Section 3.3. Section 3.3 also formulates the demand-side management problem in mircogrid level of our proposed architecture under the assumption of initial long-term energy schedule is already established. Section 3.4 derives a distributed networked control strategy to maintain the real-time microgrid stability. In Section 3.5, by considering communication delay among users and the microgrid controller, we propose a distributed networked control with prediction algorithm to achieve optimality and maintenance in the grid. Section 3.6 demonstrates the applicability and performance of the proposed methods. Section 3.7 concludes this chapter.

3.2 A Scalable, Hierarchical Control Architecture for the Grid

Much work has been done on control infrastructure for smart grid in literature. [134] provides an architecture that automates the interactions between the Utilities and their customers for demand response programs in power grid but it only supports demand response application in a centralized way without aggregation of renewables. GridStat [135] provides a flexible quality of service(QoS)-aware communication infrastructure for power grid based on internet technologies. However, it does not provide a detailed generic distributed control infrastructure. Virtual Power Plant (VPP) [136] manages aggregation of DER as a virtual conventional generation system while providing higher efficiency and more flexibility in reaction to fluctuations. VPP is a generation-level virtual architecture and requires a complicated optimization, control, and secure communication methodology.

Here we propose a scalable, hierarchical decision-making and control architecture at distribution level of smart grids. We provide a distributed control infrastructure by utilizing distributed demand-side management in microgrids through coordination and cooperation among users with RDG. Our proposed architecture not only ensures scalability in decision-making and control protocols and allows sufficient-level resource pooling. In the proposed architecture, a Utility-maintained

conventional plant and certain RDG's that are best modeled at the distribution-level feed a grid consisting of several substations (SS). Note that, wind-energy is perhaps one such RDG [62]. The objective of each substation is to maintain a nearly-flat demand-response with minimum conventional generation. Each substation is connected to N_f number of feeders each of which serves, a possibly large collection of customers. In general, the number of feeders connected to each substation may allow to be varied from substation to substation, although in practice this may not be desirable. Hence, for the brevity of notation, we assume the same fixed number N_f of feeders per substation although this can be made a variable quantity without effecting any of the development to follow. Each substation is to have energy routing capabilities, routing energy through switching, to facilitate possibility of exchanging energy among its feeders, and also be able to request power from the central plant, if needed. Therefore, we will assume that the substation has direct access and influence over conventional generation decisions. It merely ensures that at some-level in the energy-grid there is an entity with direct influence over the conventional generation decisions, and that entity will ultimately be brought to balance out the supply-demand equation. In the following, we assume that this entity's presence can be assumed at the substation level decision making.

Our goal is to ensure the model we propose for the smart-grid decision-making and interactions is scalable without precluding resource pooling capability which is one of the most important strategies to energy efficiency, maximizing energy utilization in the grid [137]. A tradeoff to stability and resource pooling can be achieved by using the notion of microgrids: The DERs served by a feeder can be divided in to several microgrids [62]. The assumption is that each microgrid is a collection of customers that are nearly self-contained and this provides a hierarchical smart-grid decision-making and control architecture as shown in Fig. 3.1 that is both scalable and allows sufficient resource pooling. We assume that there is a microgrid controller unit that is the interface between any given microgrid and the feeder. At each feeder-level there is a Master controller that interfaces with microgrid controllers of each microgrid. Master controllers interface at the substation level. In this hierarchical decision-making and control model, each microgrid is to be nearly self-contained, then each feeder is to be self-contained, and finally each substation is to be self-contained. Any changes at the lowest levels are to mostly effect the microgrid in question

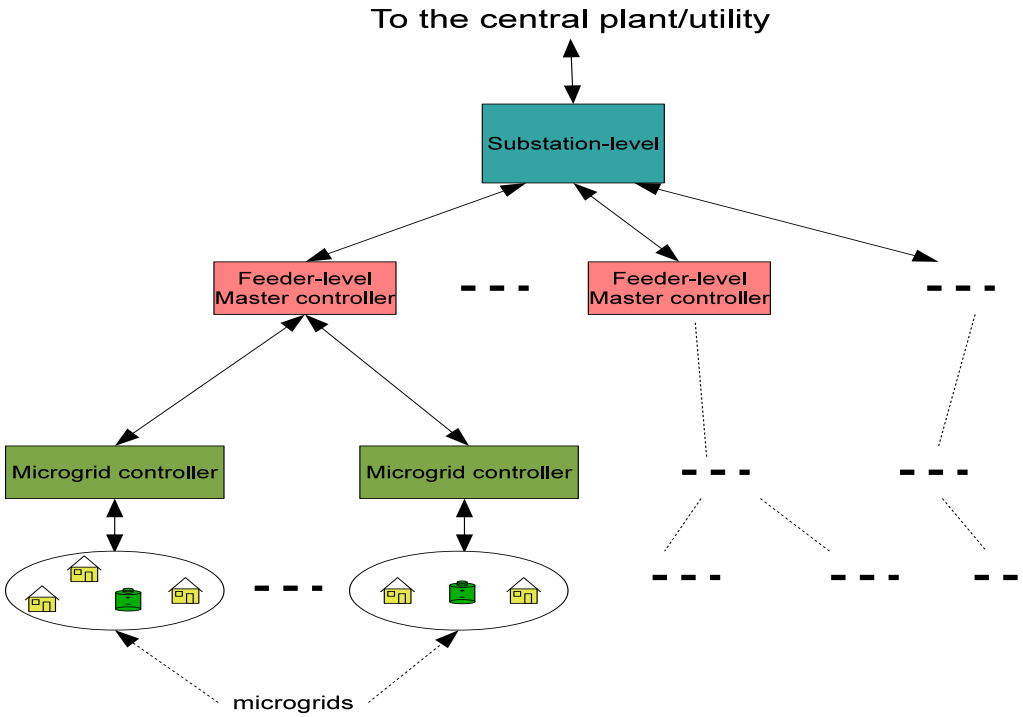


Figure 3.1: A hierarchical decision-making and control architecture for future smart-grid.

and it will only effect the Master controllers at the feeder level only if the change is too great to be resolved within the microgrid's own capabilities and resources. The same applies with respect to each feeder and substation level interaction, leading to a highly distributed and easily scalable decision-making and control architecture for the grid operation.

Within this scalable and hierarchical model, our proposed interactive framework starts with the predicted information from the customers on their look-ahead energy requirements for the next scheduling period T_s. Usually, but not necessarily, this long-term scheduling period can be of one day duration. To be specific, let us assume that at the 0-th hour of the day, each customer unit informs its predicted energy profile for the next T_s the duration (of, say, 24 hours) to its microgrid controller. This profile essentially indicates at each hour (say) of the day how much energy the customer would need from the grid or will be trading to the grid. Each microgrid controller is responsible for maintaining a nearly self-contained grid operation within its microgrid. For that,

based on all input customer energy profiles, it computes an optimized energy supply plan for the T_s duration that ensures near-constant energy draw from the feeder. The microgrid controller performs this optimization subject to various constraints on the amount of maximum expected energy from the feeder level, the energy-link capacities, fairness, QoS requirements and priorities of each customers [137]. The current framework allows systematically extending the real-time interactive smart-grid framework proposed in our earlier work [138] from microgrid level only to include the whole grid.

The long-term schedule computed by each microgrid controller is informed to all customer units as well as to the Master controller at the feeder-level. The Master controller at each feeder receives the nearly-equalized energy profiles from each microgrid under its control and performs an optimization to self-sustain its power consumption by routing positive energy flow from one microgrid to another with negative at a given time of the schedule. The computed energy schedule by the Master controller at each feeder will then be passed on to the substation level which will then perform a similar scheduling optimization among all its feeders. As can be seen from this description, the proposed architecture provides granularity at each level to ensure scalability in decision-making and control protocols as well as allow suffcient-level of resource pooling.

The above long-term or day-ahead, scheduling provides a nominal operating point for the nodes in the grid. However, these schedules are based on predicted values of loads and generations which will inevitably deviate from the actual values during the real-time operation [138]. These inevitable real-time departures from the agreed-upon schedule are to be managed within each microgrid at a finer time-scale (short-term or real-time). While there are various ways to enforce or incentivize the customers within a microgrid to pursue this collective goal of self-sustainability of a microgrid, in this chapter we limit ourselves to distributed networked control strategies: i.e. we assume that each customer unit is to observe the state, or a measurement related to the state, of the microgrid and to implement a classical control on its own DER to help maintain the required grid performance.

The focus of the rest of this chapter, is the demand-side management within each microgrid at a finer time scale to ensure that the state of the grid is maintained at a scheduled operating point. In

other words, we do not consider initial long-term scheduling problem within microgrid, feeder and substation levels but rather assume that these are already established (see, for example, [67, 138] for possible approaches to solve the initial scheduling problem).

3.3 Demand-side management within a Microgrid

Consider the following short-term demand-side management problem within a microgrid in grid-connected mode[1], which consists of N customers and one microgrid controller P as in Fig. 3.2. Each customer has energy storage and is equipped with RDG. At the beginning of each long-term day-ahead scheduling period T_s, each customer unit informs its predicted energy profile for the next T_s duration to its microgrid controller P over the communication channel $\mathbf{CH}_i$. Based on all input customer energy profiles, microgrid controller P computes an optimized demand response schedule for the T_s duration that ensures near-constant energy draw from the feeder-level. Then, microgrid controller P informs this long-term demand response schedule to all customer units over feedback communication channels $\mathbf{Ch}_i$ in its microgrid at the beginning of the scheduling period T_s, say at time step $k = 0$.

Let $\mathbf{Z}(k)$ represent the scheduled operational point at microgrid controller P at time step k according to the established long-term schedule of duration T_s where $k = 0, 1, 2 \cdots, T_s$. Let $\mathbf{y}_i(k)$ and $\mathbf{y}(k)$ denote user i's real-time net power consumption and the net power consumption over all users at time step k. Since the schedules are based on predicted values of loads and generations, the actual values of $\mathbf{y}(k)$ during the real-time operation will inevitably deviate from the agreed-upon schedule $\mathbf{Z}(k)$ which is be managed within each microgrid in real-time. This deviation is described by power offset $\mathbf{e}(k)$, which is defined as $\mathbf{e}(k) = \mathbf{Z}(k) - \sum_{i=1}^{N} \mathbf{y}_i(k) \quad \text{for} \quad 0 \leq k \leq T_s$. The goal is that with the support of energy storage and RDG, each customer unit observes the state of the microgrid and adjusts its net power consumption $\mathbf{y}_i(k)$ to maintain load-supply balance within the

[1]The management goals are quite different in two operation modes. While in grid-connected mode, the controlling power is a necessary management function rather than the frequency and voltage in islanded mode [62]. In this paper, we only consider microgrid in grid-connected mode.

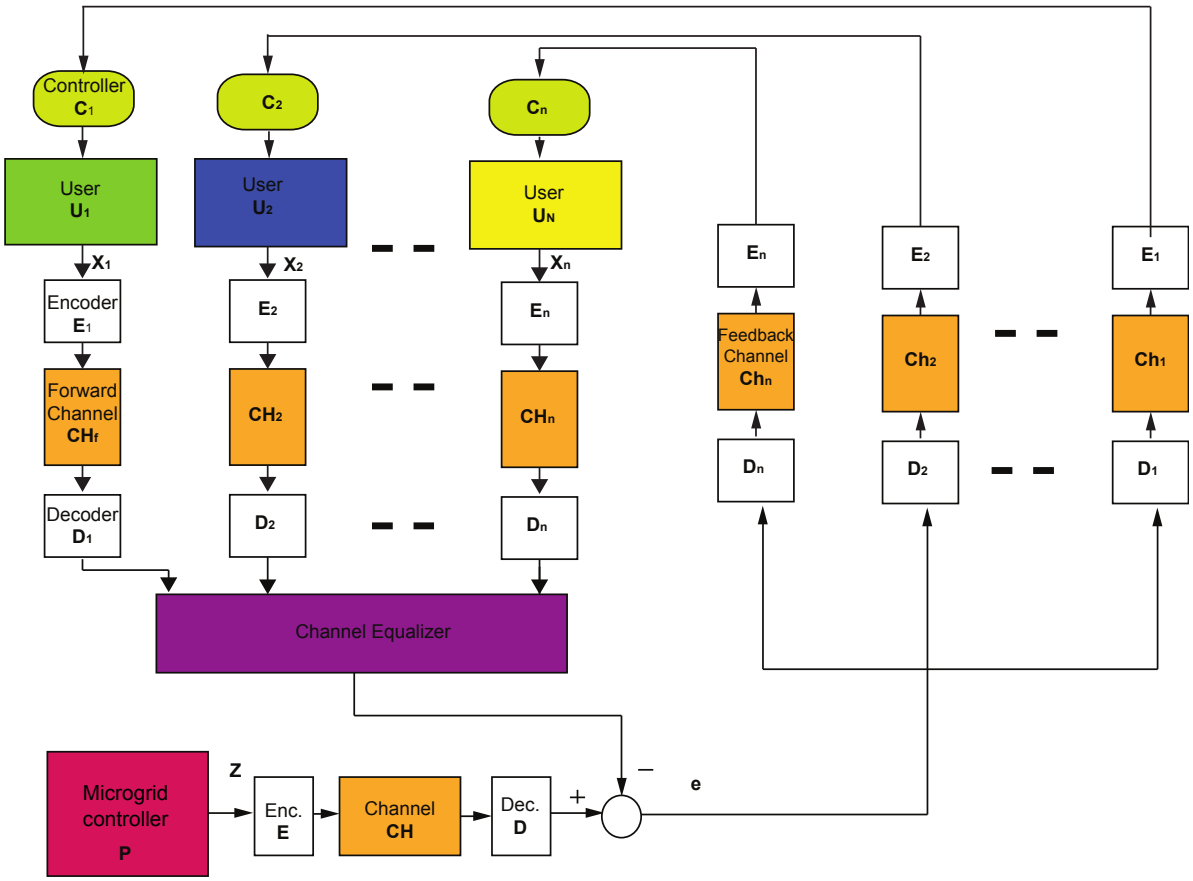


Figure 3.2: Microgrid with N users and one microgrid controller P.

microgrid.

3.3.1 User Model

Consider user i as an agent equipped with RDG and energy storage devices and with decision-making capability to decide its net power consumption $\mathbf{y}_i(k)$ from microgrid based on load-supply information. User i's load demand can be supplied by RDG, energy storage and microgrid. Renewable sources such as wind and solar, are usually intermittent, uncertain, and uncontrollable, which results in fluctuations in RDG energy supply. On the other hand, the user load demand has very limited freedom to be controlled through load shaping because of non-elastic loads [139]. However, energy storage serves as an compensation to RDG making user load demand from a mi-

crogrid more elastic. User i can store excess renewable energy generation into energy storage and supply its load with energy storage instead of requesting energy from microgrid when renewable energy is insufficient[2]. In this way, by using energy storage, user i can supply its load demand and also be able to adjust its net power consumption $\mathbf{y}_i(k)$ from microgrid such that the aggregate net power consumption $\mathbf{y}(k)$ can match to the agreed-on energy schedule $\mathbf{Z}(k)$.

Upon receiving load-balance information $\mathbf{e}(k)$ from microgrid controller P, user i determines its net power consumption $\mathbf{y}_i(k)$. To take into account the stochastic characteristics of RDG and user load, we formulate this decision-making process as a stochastic system as follows [141]:

$$\begin{aligned} \mathbf{x}_i(k+1) &= \mathcal{F}_i(\mathbf{x}_i(k), \mathbf{u}_i(k), \mathbf{w}_i(k)), \\ \mathbf{y}_i(k) &= \mathcal{G}_i(\mathbf{x}_i(k)), \end{aligned} \tag{3.1}$$

where $\mathbf{x}_i(k)$ is the state of user i at time k, $\mathbf{w}_i(k)$ is random disturbance, i.e. fluctuation in RDG energy supply, disturbance from user load, $\mathbf{u}_i(k)$ is user i's control decision, $\mathbf{y}_i(k)$ is user i's net power consumption at time k and $\mathcal{F}_i(\cdot)$, $\mathcal{G}_i(\cdot)$ are deterministic functions. $\mathcal{F}_i(\cdot)$ describes user i's state transition dynamics. $\mathcal{G}_i(\cdot)$ describes the relation between user i's state and its net power consumption.

In this formulation, at time k, user i has its current state $\mathbf{x}_i(k)$ related to its net power consumption $\mathbf{y}_i(k)$. Upon receiving power offset $\mathbf{e}(k)$ from microgrid controller, user i makes control decision $\mathbf{u}_i(k)$ on its net energy usage at the next time step. With control decision $\mathbf{u}_i(k)$, and state $\mathbf{x}_i(k)$, user i applies function $\mathcal{F}_i(\cdot)$ to update its state $\mathbf{x}_i(k+1)$ and outputs its net power usage $\mathbf{y}_i(k+1)$ using function $\mathcal{G}_i(\cdot)$. This stochastic system in (3.1) describes the decision-making process of each user and also takes into account the stochastic characteristics from renewable energy supply and user load consumption by considering them as random disturbance $\mathbf{w}_i(k)$.

In (3.1), functions $\mathcal{F}_i(\cdot)$ and $\mathcal{G}_i(\cdot)$ could be nonlinear and complex functions for which analytical results may be not available [142]. In order to derive useful mathematical results, we only

[2]Energy storage devices have limited capacites [140]. Here we assume that the short-term user energy consumption is within this limit.

consider linear functions $\mathcal{F}_i(\cdot)$ and $\mathcal{G}_i(\cdot)$. Hence, we can simplify (3.1) into a linear time-invariant (LTI) stochastic system [141] as follows:

$$\begin{aligned} \mathbf{x}_i(k+1) &= A\mathbf{x}_i(k) + B\mathbf{u}_i(k) + \mathbf{w}_i(k), \\ \mathbf{y}_i(k) &= C\mathbf{x}_i(k) \quad \text{for} \quad 0 \leq k \leq T_s, \end{aligned} \tag{3.2}$$

where system state $\mathbf{x}_i(k)$ and net load $\mathbf{y}_i(k)$ take values in $\mathbb{R}^p$ and control decision $\mathbf{u}_i(k)$ takes values in $\mathbb{R}^q$ for $p, q \in \{1, 2, \cdots\}$. The initial state $\mathbf{x}_i(0)$ and noises $\mathbf{w}_i$ are assumed to be zero mean Gaussian random variables with covariance matrices Σ_x and Σ_w. Here we assume that the system is observable and controllable such that each customer unit can observe the state $\mathbf{x}_i(k)$ and implement control decision $\mathbf{u}_i(k)$ to adjust net load $\mathbf{y}_i(k)$. Our objective is to design a control strategy for each user i such that energy utilization in the smart grid is optimized, which is characterized in term of the second moment of power offset $\mathbf{e}(k)$. In the following sections, we consider distributed networked control strategy based on the assumption that information exchange is feasible among users.

3.4 Distributed Networked Control Strategy

As mentioned in Section 1.2.2, optimization techniques provide a system-level analysis but operating in a centralized manner, while game theory provides distributed decision-making strategies but lacking a system-level design. To obtain a system-level optimization of the demand-side management problem while providing distributed analytical techniques to microgrids with DG, we propose a distributed control algorithm with information exchange among users. Under the assumption of cooperative users and inter-user communication, the proposed approach applies optimal control to aggregate user energy usage within a microgrid and it could be implemented distributedly in each RDG-equipped user. A tradeoff can be made between control effort and tracking accuracy. This approach requires digital communication infrastructure among users to enable information exchange. When communication delay exists, it also requires previous control inputs of all users, meaning that each node should have data storage capability.

The demand-side management problem we posed in Section 3.3 is to match the aggregate user energy consumption $\sum_i^N \mathbf{y}_i(k)$ with the agreed-on schedule $\mathbf{Z}(k)$ so that energy utilization is optimized in the smart grid, which is characterized by the mean squared error of power offset $\mathbf{e}(k)$ given by $\mathbb{E}[\mathbf{e}(k)^T\mathbf{e}(k)]$. This is a tracking control problem in optimal control theory [143]. Indeed, from optimal control theory, we could define a quadratic cost function $J_0^{T_s}$ in terms of tracking accuracy and control effort relating to the user power utilization dynamics in (3.2) for $0 \leq k \leq T_s$ as follows:

$$J_0^{T_s} = \mathbb{E}\Big[\frac{1}{2}\sum_{k=0}^{T_s-1}(\mathbf{e}(k)^T Q\mathbf{e}(k) + \sum_{i=1}^{N}\mathbf{u}_i(k)^T R\mathbf{u}_i(k)) + \frac{1}{2}\mathbf{e}(T_s)^T F\mathbf{e}(T_s)\Big], \tag{3.3}$$

where F and Q are $p \times p$ symmetric, positive semidefinite matrices, and R is a $q \times q$ symmetric, positive definite matrix. The first term $\mathbf{e}(k)^T Q\mathbf{e}(k)$ is a measure of tracking accuracy between the total user net consumption $\sum_{i=1}^N \mathbf{y}_i(k)$ and agreed-upon energy schedule $\mathbf{Z}(k)$. The term $\sum_{i=1}^N \mathbf{u}_i(k)^T R\mathbf{u}_i(k)$ represents the control effort over all customers and $\mathbf{e}(T_s)^T F\mathbf{e}(T_s)$ represents the terminal tracking accuracy. Hence, we have the optimization problem in the following form:

$$\begin{aligned}
\text{minimize} \quad & J_0^{T_s} \\
\text{subject to} \quad & \mathbf{x}_i(k+1) = A\mathbf{x}_i(k) + B\mathbf{u}_i(k) + \mathbf{w}_i(k), \\
& \mathbf{y}_i(k) = C\mathbf{x}_i(k),\ \mathbf{x}_i(0) = \mathbf{x}_{i,0}, \\
& Q, F \geq 0, \quad R > 0,
\end{aligned} \tag{3.4}$$

In order to achieve optimal energy utilization, we need to minimize the cost function to acquire an optimal control input $\mathbf{u}_i(k)$ for each user.

3.4.1 Derivation of the Distributed Optimal Controller:

The optimization problem defined in (3.4) is an additive disturbance stochastic LQR problem [152] with disturbance $\mathbf{w}_i(k)$ as an additive white Gaussian noise. According to certainty-equivalence principle [152], this additive disturbance stochastic problem can be solved as a deterministic problem with stochastic variable $\mathbf{w}_i(k)$ replaced by its average value $\mathbb{E}[\mathbf{w}_i(k)] = 0$. Thus, the optimal

control for this additive disturbance stochastic problem is identical to the optimal control for the deterministic case as follows:

$$\begin{aligned}
\text{minimize} \quad & \tilde{J}_0^{T_s} = \frac{1}{2}\sum_{k=0}^{T_s-1}\big(\mathbf{e}(k)^T Q\mathbf{e}(k) + \sum_{i=1}^{N}\mathbf{u}_i(k)^T R\mathbf{u}_i(k)\big) + \frac{1}{2}\mathbf{e}(T_s)^T F\mathbf{e}(T_s) \\
\text{subject to} \quad & \mathbf{x}_i(k+1) = A\mathbf{x}_i(k) + B\mathbf{u}_i(k), \\
& \mathbf{y}_i(k) = C\mathbf{x}_i(k),\ \mathbf{x}_i(0) = \mathbf{x}_{i,0}, \\
& Q, F \geq 0, \quad R > 0.
\end{aligned} \tag{3.5}$$

In the remaining of this section, instead of solving the stochastic problem in (3.4), we will derive the optimal controller for the deterministic case in (3.5)[3]. First, we define the Hamiltonian for the optimization problem in (3.5) as

$$\begin{aligned}
&\mathcal{H}\big(\mathbf{x}_i(k), \mathbf{u}_i(k), \boldsymbol{\gamma}_i(k+1)\big) \\
&= \frac{1}{2}\sum_{k=0}^{T_s-1}\big(\mathbf{e}(k)^T Q\mathbf{e}(k) + \sum_{i=1}^{N}\mathbf{u}_i(k)^T R\mathbf{u}_i(k)\big) + \sum_{i=1}^{N}\boldsymbol{\gamma}_i^T(k+1)[A\mathbf{x}_i(k) + B\mathbf{u}_i(k)],
\end{aligned} \tag{3.6}$$

where $\boldsymbol{\gamma}_i^T(k+1)$ for $1 \leq i \leq N$ are Lagrange multipliers to be determined. The Hamiltonian is obtained by adjoining the original cost function (3.3) with plant relation (3.2) using Lagrange multiplier $\boldsymbol{\gamma}_i^T(k+1)$. Define $E = BR^{-1}B^T$, $V = C^T QC$ and $W = CQ$. From calculus of variations [144], the Euler-Lagrange difference equation corresponding to the Hamiltonian (3.6) with variables $\boldsymbol{\gamma}_i(k+1)$, $\mathbf{x}_i(k)$ and $\mathbf{u}_i(k)$ leads to the following required conditions for an extremum:

$$\mathbf{x}_i(k+1) = \frac{\partial\mathcal{H}}{\partial\boldsymbol{\gamma}_i(k+1)} = A\mathbf{x}_i(k) + B\mathbf{u}_i(k), \tag{3.7}$$

$$\boldsymbol{\gamma}_i(k) = \frac{\partial\mathcal{H}}{\partial\mathbf{x}_i(k)} = A^T\boldsymbol{\gamma}_i(k+1) + V\sum_{i=1}^{N}\mathbf{x}_i(k) - W\mathbf{Z}(k), \tag{3.8}$$

$$0 = \frac{\partial\mathcal{H}}{\partial\mathbf{u}_i(k)} = R\mathbf{u}_i(k) + B^T\boldsymbol{\gamma}_i(k+1). \tag{3.9}$$

[3]The distributed optimal controller in (3.5) is an extension of the standard optimal controller [143], because its optimization problem is defined over N different systems instead of one single system.

From the Eular-Lagrange difference equation and the established long-term schedule of duration T_s, we have the following boundary condition at time $k = T_s$:

$$\begin{aligned}\boldsymbol{\gamma}_i(T_s) &= \frac{\partial}{\partial \boldsymbol{x}_i(T_s)}\big[\frac{1}{2}\mathbf{e}(T_s)^T F\mathbf{e}(T_s)\big] \\ &= C^T FC\sum_{i=1}^{N}\mathbf{x}_i(T_s) - C^T F\mathbf{Z}(T_s),\end{aligned} \tag{3.10}$$

From (3.10), it is clear that $\boldsymbol{\gamma}_i(T_s) = \boldsymbol{\gamma}_j(T_s)$ for $1 \leq i, j \leq N$. By substituting (3.10) into (3.8), we have that $\boldsymbol{\gamma}_i(k) = \boldsymbol{\gamma}_j(k)$ for $1 \leq i, j \leq N$ and $0 \leq k \leq T_s$, which together with (3.9) and (3.10) yield

$$\mathbf{u}_i(k) = \mathbf{u}_j(k) = -R^{-1}B^T\boldsymbol{\gamma}_i(k+1), \tag{3.11}$$

for $1 \leq i, j \leq N$ and $0 \leq k \leq T_s$. From the nature of the boundary condition (3.10), we may assume a transformation as follows

$$\boldsymbol{\gamma}_i(k) = P(k)\sum_{i=1}^{N}\mathbf{x}_i(k) - g(k), \tag{3.12}$$

where matrix $P(k)$ and vector coefficient $g(k)$ are needed to be determined. Substituting (3.12) into (3.7) and solving for $\mathbf{x}_i(k)$ yields that

$$\mathbf{x}_i(k) = A^{-T}[EP(k+1)\sum_{i=1}^{N}\mathbf{x}_i(k+1) + \mathbf{x}_i(k) - Eg(k+1)]. \tag{3.13}$$

Using (3.13) and (3.12) in equation (3.8), we have

$$\begin{aligned}&[(V - P(k))A^{-1}(NEP(k+1)+1)A^T P(k+1)]\sum_{i=1}^{N}\mathbf{x}_i(k+1) \\ &+ g(k) - [A^T + N(V - P(k))A^{-1}E]g(k+1) - W\mathbf{Z}(k) = 0.\end{aligned} \tag{3.14}$$

Since (3.14) holds for all values of state $\mathbf{x}_i(k+1)$, it implies that the coefficient of $\sum_{i=1}^{N}\mathbf{x}_i(k+1)$ and the rest of the terms must vanish individually. Hence, by letting these terms in (3.14) equal to

zero, we obtain the Riccati function $P(k)$ and vector coefficient $g(k)$ as

$$P(k) = V + A^T(P^{-1}(k+1) + NE)^{-1}A, \tag{3.15}$$

$$g(k) = [A^T + N(V - P(k))A^{-1}E]g(k+1) + W\mathbf{Z}(k), \tag{3.16}$$

Comparing boundary condition (3.10) and (3.12), we have that $P(T_s) = C^T FC$ and $g(T_s) = C^T F\mathbf{Z}(K)$. Substituting (3.15), (3.16) and transformation (3.12) into (3.11), we have the controller $\mathbf{u}_i(k)$ as following:

$$\mathbf{u}_i(k) = -L(k)\sum_{i=1}^{N}\mathbf{x}_i(k) + L_g(k)g(k+1), \tag{3.17}$$

$$L_g(k) = [R + NB^T P(k+1)B]^{-1}B^T, \tag{3.18}$$

$$L(k) = L_g(k)P(k+1)A, \tag{3.19}$$

where $P(k)$ and $g(k)$ are defined in (3.15) and (3.16) with final condition $P(T_s) = C^T FC$ and $g(T_s) = C^T F\mathbf{Z}(T_s)$.

Remark 3.4.1. *In equation (3.17), the ith user's control input $\boldsymbol{u}_i(k)$ depends on not just its own current state $\boldsymbol{x}_i(k)$ but also the current states of other users. It requires information exchange among users or microgrid control broadcasting information to users such that each user could obtain the sum of states over all users in the microgrid. The feedback gain $L(k)$ and feed-forward gain $L_g(k)$ in equations (3.19) and (3.18) depend on the Riccati function $P(k+1)$ in the next time step. The Riccati function $P(k)$ could be obtained by solving backwards the discrete-time Riccati equation using the final condition $P(T_s) = C^T FC$. $\mathbf{Z}(k)$ could be obtained from power offset $\boldsymbol{e}(k)$ and aggregate user net load $\sum_{i=1}^{N}\boldsymbol{y}_i(k)$.*

3.5 Demand-side Management within Microgrid with Communication Delays

3.5.1 Same Communication Delay

Consider the demand-side management in microgrid with communication delays as in Fig. 3.2, where the feedback channels Ch_is between users and microgrid controller P are assumed time-delayed. Here we are interested in designing distributed networked control with prediction to control this microgrid with time-delayed feedback.

To illustrate that, we first start with the case that all feedback channels Ch_is have the same communication delay h. Since feedback channels have time delay, the power offset received by user i at time k is $\mathbf{e}(k-h)$. We assume that before receiving feedback information $\mathbf{e}(k)$ no control effort is applied. Hence, we have the control input of each system is time delayed $\mathbf{u}_i(k-h)$, where $\mathbf{u}_i(k)=0$ for $0\leq k\leq h$. By using reduction method [145], we transform this control system with input time delayed into a new system without delay. From the user dynamics in (3.2), the system state $\mathbf{x}_i(k+h)$ is

$$\begin{aligned}\mathbf{x}_i(k+h) &= A^h\mathbf{x}_i(k)+\sum_{l=k}^{k+h}A^{k+h-l}\big(B\mathbf{u}_i(l-h)+\mathbf{w}_i(l-h)\big)\\ &= A^h\xi_i(k),\end{aligned} \tag{3.20}$$

where new system state $\xi_i(k)=\mathbf{x}_i(k)+\sum_{l=k}^{k+h}A^{k-l}\big(B\mathbf{u}_i(l-h)+\mathbf{w}_i(l-h)\big)$. Then we have the new system described as follows:

$$\begin{aligned}\xi_i(k+1) &= A\xi_i(k)+A^{-h}B\mathbf{u}_i(k)+A^{-h}\mathbf{w}_i(k),\\ \eta_i(k) &= C\xi_i(k),\end{aligned} \tag{3.21}$$

where $\eta_i(k)$ is the system output.

For this new system, assume that $\mathbf{u}_i(k)=0$ for $0\leq k\leq h$. It is reasonable that we ignore the cost term related to $\mathbf{w}_i(k)$ for $0\leq k\leq h$, because the additive noise is not controllable. Then the

cost function in (3.3) becomes

$$J_0^{T_s} = \frac{1}{2}\sum_{k=0}^{h} \mathbf{e}(k)^T Q \mathbf{e}(k) + J_h^{T_s}, \tag{3.22}$$

where $J_h^{T_s}$ is the cost function from time h to T_s. The first term in (3.22) is constant. Hence, to minimize $J_0^{T_s}$, we need to minimize the term $J_h^{T_s}$, which could be written as

$$J_h^{T_s} = \frac{1}{2}\sum_{k=0}^{T_s-h} \left[\alpha(k)^T F \alpha(k) + \sum_{i=1}^{N} \mathbf{u}_i(k)^T R \mathbf{u}_i(k)\right] + \frac{1}{2}\beta(T_s)^T F \beta(T_s), \tag{3.23}$$

where $\alpha(k) = \sum_i CA^h\xi_i(k) - \mathbf{Z}(k+h)$ and $\beta(T_s) = \sum_i CA^h\xi_i(T_s-h) - \mathbf{Z}(T_s)$. Following the derivation in Section 3.4, we could find the following optimal controller

$$\begin{aligned}
\mathbf{u}_i(k) &= -L(k)\sum_i \left(\mathbf{x}_i(k) - \overline{\mathbf{u}}(k)\right) + L_g(k)g(k+1),\\
\overline{\mathbf{u}}(k) &= \sum_{l=k}^{k+h-1} A^{k-1-l}B\sum_{i=1}^{N}\mathbf{u}_i(l-h),
\end{aligned} \tag{3.24}$$

where the parameters is defined as follows

$$\begin{aligned}
P(k) &= A^T P(k+1)[I + N\overline{E}P(k+1)]^{-1}A + \overline{V},\\
g(k) &= [A^T + N(\overline{V} - P(k))A^{-1}\overline{E}]g(k+1) + W\mathbf{Z}(k+h),\\
L_g(k) &= [R + NB^T A^{-h\prime}P(k+1)A^{-h}B]^{-1}B^T A^{-h\prime},\\
L(k) &= L_g(k)P(k+1)A,
\end{aligned} \tag{3.25}$$

where $\overline{V} = C^T Q C A^h$, $\overline{E} = A^{-h}BR^{-1}B^T A^{-h\prime}$, $W = C^T F$, $P(T_s) = C^T F C A^h$, $g(T_s) = C^T F \mathbf{Z}(T_s+h)$ and $A^{-h\prime}$ denotes $\left(A^T\right)^{-h}$. From (3.24), we know that for each user, the system input $\mathbf{u}_i(k)$ at time k depends on the previous inputs of all users $\sum_{i=1}^{N}\mathbf{u}_i(l)$ for $0 \le l \le k$. Hence, we need each user to communicate with other users not just its current state $\mathbf{x}_i(k)$ but also its control input $\mathbf{u}_i$ or microgrid controller to broadcast this information to users.

3.5.2 Different Communication Delays

Next, we extend the demand-side management within microgrid over the same time delay to the case with different time delays. Let us denote the pure time delay for each feedback channel as h_i,

where $h_i \neq h_j$, if $i \neq j$. Without loss of generality, we assume that $h_1 \leq h_2 \leq \cdots \leq h_N$ and the system input $\mathbf{u}_i(k) = 0$ for $0 \leq k \leq h_i$. The cost function in (3.3) can be written as

$$J_0^{T_s} = J_0^{h_1} + J_{h_1}^{h_2} + J_{h_2}^{h^3} + \cdots + J_{h_N}^{T_s}, \tag{3.26}$$

where $J_0^{h_1}$ is constant and we need to minimize $J_{h_1}^{h_2}, \cdots, J_{h_N}^{T_s}$ in order to minimize $J_0^{T_s}$. Then, for each J_a^b, we can optimize it based on the number of users have non-zero system input $\mathbf{u}_i(k) \neq 0$ for $a \leq k \leq b$. For each time slot $h_p \leq k \leq h_q$, we can find the optimal control input $\mathbf{u}_i(k)$ for user $1 \leq i \leq h_q$ by following the derivation in the same communication delay case in Section 3.5.1.

3.6 Simulated Performance Examples

In this section, we will demonstrate the numerical performance and applicability of the proposed approaches.

First, we show the performance of the above distributed networked control strategy in a microgrid problem. In this simulation, we consider the following system parameters. The following parameters do not take realistic values, but they are chosen for simplicity and to better demonstrate the effectiveness of the proposed approach. $A = \begin{bmatrix} 0.8907 & 0.8704 & -0.0785 \\ -0.2019 & 0.7166 & -0.1449 \\ -0.1784 & -0.0873 & 0.8659 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0.1000 & 0 & 0.15 \end{bmatrix}$, $N = 20$, $l = 3$, $K = 60$, $F = Q = I$, $R = I$, $\Sigma_w = 0.01I$ and $\Sigma_x = I$ [4]. We assume that the established energy schedule $Z(k)$ is modeled as follows

$$\mathbf{Z}(k) = C_p \mathbf{X}(k), \;\; \mathbf{X}(k+1) = A_p \mathbf{X}(k) + \mathbf{n}_p(k), \tag{3.27}$$

where $\mathbf{n}_p(k)$ is a zero mean Gaussian random variable with covariance matrix $\Sigma_p = 0.04I$, $A_p =$

[4]Note that, the elements of matrix A and A_p are chosen such that the largest eigenvalues of A and A_p equal to one, while the systems remain marginally stable.

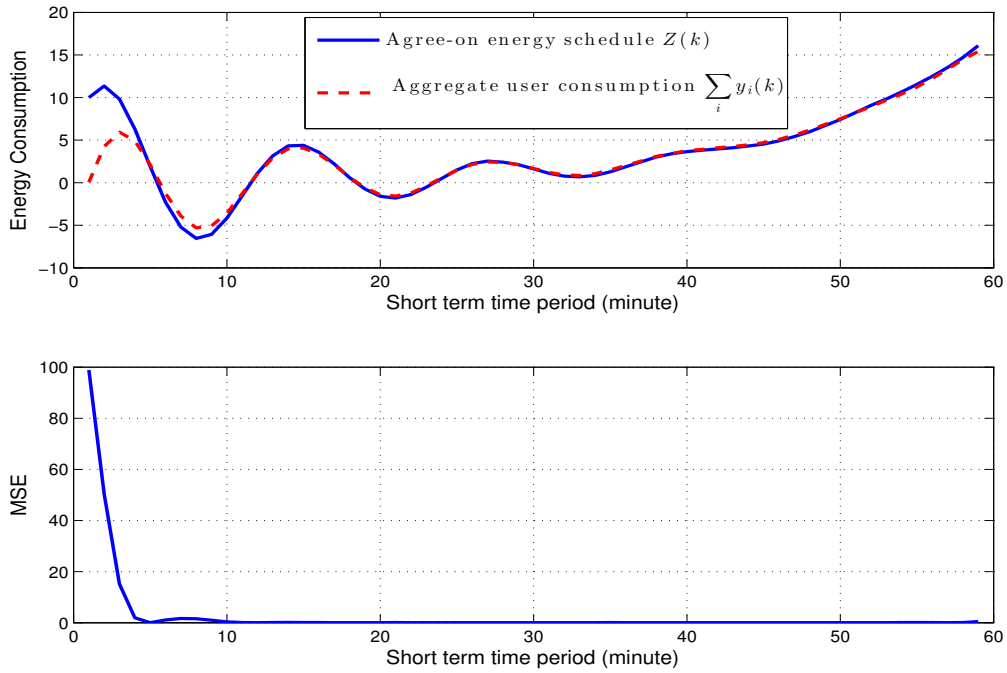


Figure 3.3: Ditributed networked control with prediction for microgrid with multiple time delays – (top) short term comparison between total user net consumption $\mathbf{y}(k)$ and energy schedule $\mathbf{Z}(k)$, (down) mean squared error.

$$\begin{bmatrix} 0.9763 & 0.9541 & -0.0861 \\ -0.2213 & 0.7855 & -0.1588 \\ -0.1956 & -0.0957 & 0.9492 \end{bmatrix} \text{ and } C_p = \begin{bmatrix} 0.1000 & 0 & 0.15 \end{bmatrix}.$$

Figure 3.3 shows the short-term minutely energy schedule $\mathbf{Z}(k)$ and aggregate user consumption $\sum_{i=1}^{N} \mathbf{y}_i(k)$ for distributed control with inter-user communications. It can be seen from Fig. 3.3 that the proposed approach drives the aggregate user consumption $\sum_{i=1}^{N} \mathbf{y}_i(k)$ to track the agreed-on energy schedule $\mathbf{Z}(k)$ within 10 minutes and the MSE $\mathbb{E}[\mathbf{e}(k)^T \mathbf{e}(k)]$ decreases as time k increases. Figure 3.4 shows the average MSE over time $\frac{1}{T_s} \sum_k \mathbb{E}[\mathbf{e}(k)^T \mathbf{e}(k)]$ as a function of number of users in microgrid with same initial condition $\Sigma_x = \mathbf{0}$ and $\Sigma_w = \mathbf{0}$. It can be seen that the average MSE decreases as the number of users N increases, which is advantageous and beneficial. It implies that by using distributed control approach, better energy schedule maintenance in microgrid can be achieved as more users participate and take cooperative coordination in adjusting

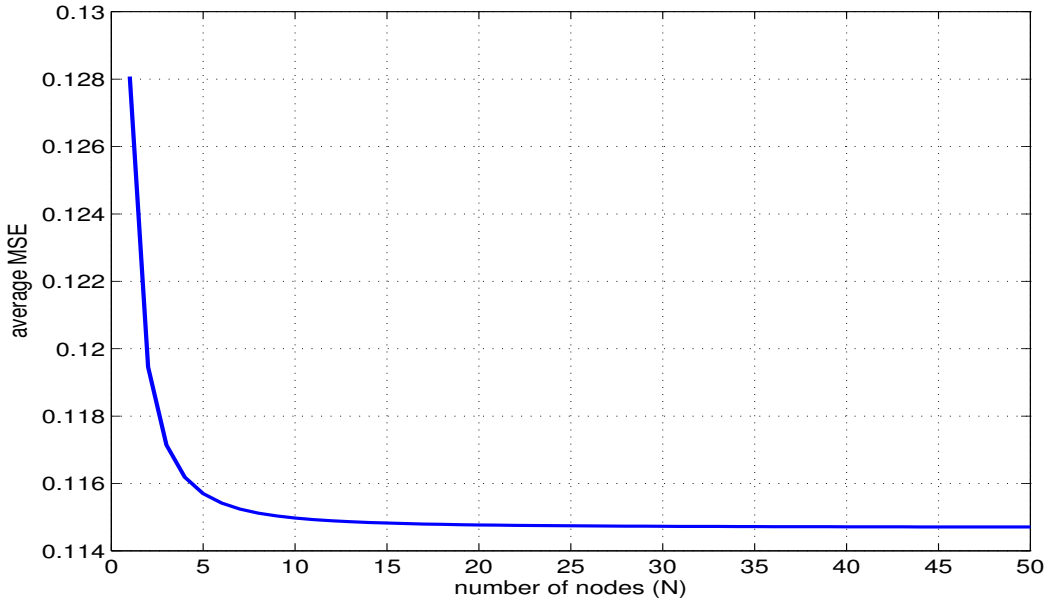


Figure 3.4: Performance versus number of users in microgrid – average mean squared error over number of users N (distributed control approach).

energy consumption, which is not the case for non-cooperative game theory approach.

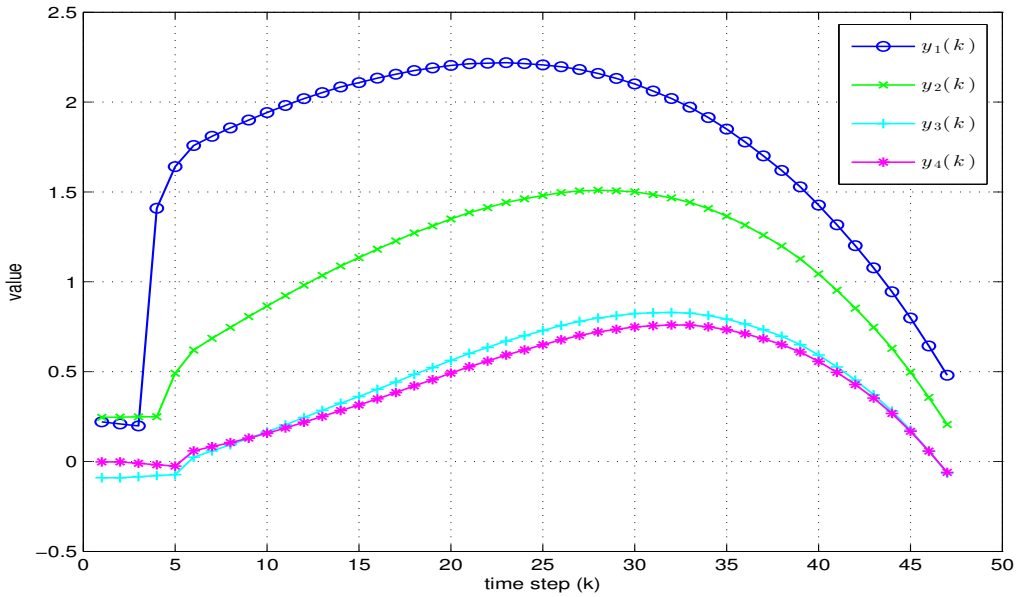


Figure 3.5: Ditributed networked control with prediction for microgrid with multiple time delays – user net power consumption $\mathbf{y}_i(k)$.

Next, we demonstrate the performance of the proposed algorithm on the demand-side management in microgrid with time-delayed feedback. The established energy schedule $\mathbf{Z}(k)$ defined in (3.27) with the same parameters except for $A_p = \begin{bmatrix} 1.0536 & 0.1044 & -0.0009 \\ -0.0259 & 1.0327 & -0.0187 \\ -0.0209 & -0.0010 & 1.0392 \end{bmatrix}$. Time delays $h_1 = 1$, $h_2 = 2$, $h_3 = h_4 = 3$ and $F = Q = 100I$, $R = 0.01I$. The rest parameters are defined as the same as in the second simulation. Fig. 3.5 shows short-term minutely each user's net power consumption $\mathbf{y}_i(k)$. It can be seen that each user shows different response latencies due to time delays in reception of power offset $\mathbf{e}(k)$. Also, because of time delays and different initial conditions $\mathbf{x}_i(0)$, the users' outputs are different from one another but show the same pattern. Compared to previously proposed optimization approaches that requires collective information at centralized controller, our approach is distributed, fully scalable and easy to implement in each user to achieve the optimal performance. In Fig. 3.6, it could be seen that the MSE starts increasing at the beginning due to no feedback input. After time delay h_1, the total user net consumption $\mathbf{y}(k)$ starts tracking and follows the energy schedule $\mathbf{Z}(k)$ and the MSE decreases. The system could respond as fast as the smallest delay h_1 instead of waiting till the largest delay h_4 as shown in Section 3.5.2.

(optimality principle)

3.7 Conclusions

In this chapter, we proposed a hierarchical decision-making and control architecture for smart grid in which distributed customers equipped with RDG interact and trade energy in the grid. Within this framework, we proposed a distributed networked control strategy with prediction to solve the demand-side management problem encountered within a microgrid with time delay. Our approaches are distributed, fully scalable and easy to implement, which provides nearly perfect performance with the cost of communication. However, here we do not consider factors such as pricing information, storage cost and utilities of the adjustable loads, etc. Possible further work

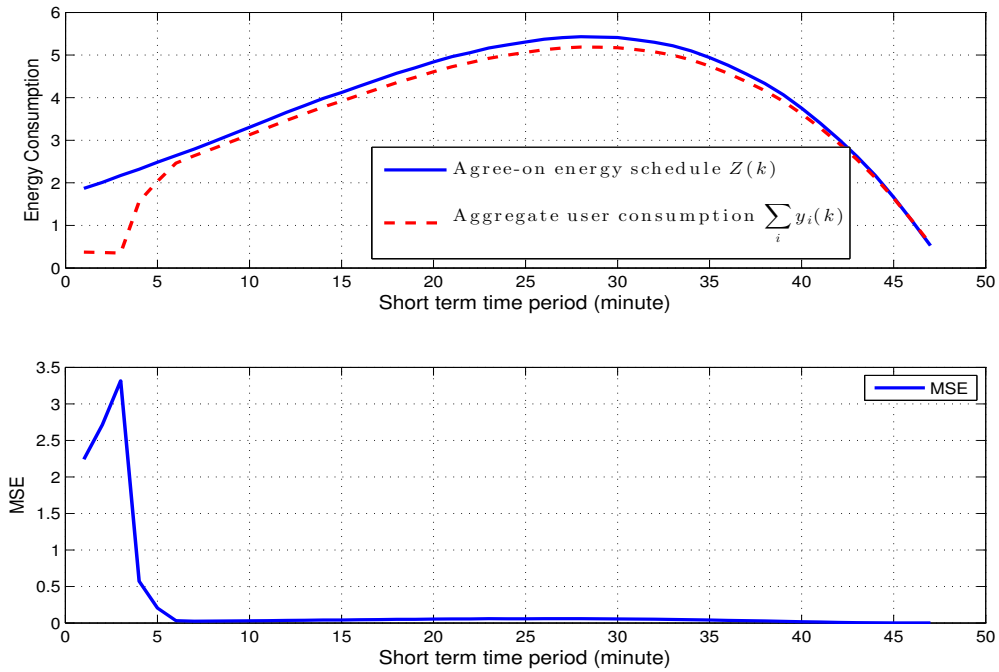


Figure 3.6: Ditributed networked control with prediction for microgrid with multiple time delays – (top) comparison between total user net consumption $\mathbf{y}(k)$ and energy schedule $\mathbf{Z}(k)$, (down) mean squared error.

involves relaxing the assumption of information exchange among users and designing a distributed control method under that assumption.

In the next chapter, we discuss the application of graph theory and cooperative control in another type of networked multi-agent system – automated highway system. We consider the string stability problem in vehicle platoons with an inter-vehicle communications network, in which each vehicle tries to maintain a fixed distance from its predecessor. We propose an leader-following consensus protocol and derive the sufficient conditions, in terms of communication topology and control parameters for string stability in such system.

Chapter 4

Disturbance Attenuation and Stabilization in Automated Highway Systems

4.1 Introduction

In this chapter, we consider the problem of controlling a platoon of vehicles in which each vehicle tries to maintain a fixed distance from its predecessor, which is an instance of the so-called "string stability" problem. Our model includes a communication network among the platoon of vehicles and an associated leader-following consensus protocol. We derive the sufficient conditions, in terms of communication topology and control parameters, for string stability by using the leader-following consensus algorithm. Comparison of our proposed algorithm with previously proposed control methods in literature shows that ours is a generalization that encompasses previous approaches as special cases obtained by assuming specific communication topologies. Simulation results and performance in terms of disturbance propagation are also given, showing that the proposed leader-following consensus protocol leads to almost the same level of performance as the previous approaches, while the proposed algorithm has additional advantages of scalability, robustness and fully distributed implementation.

The rest of the chapter is organized as follows: Section 4.2 formulates the string stability problem in a vehicle platoon and describes the assumed system model of each vehicle and network model. In Section 4.3, we propose the leader-following consensus algorithm over a communication network and derive the sufficient conditions for achieving string stability. We also point out previously proposed control approaches found in literature as special cases of our proposed method with specific types of communication graphs. Section 4.4 shows the simulation results and compares the performance of our proposed method with that of previously proposed control approaches. The conclusions are provided in Section 4.5.

4.2 Problem Formation

4.2.1 System Model

Consider a string of $N+1$ systems as a platoon as in Fig. 4.1. Let $x_0(t)$ denote the position of the leader vehicle and $x_i(t)$ $(1 \leq i \leq N)$ denote the position of the ith follower vehicle in the string. Assume that all the systems have the same model and are controlled by identical control laws. The desired spacing δ between each two consecutive vehicles is a constant. We assume that the string of vehicles start with zero spacing errors. The lead vehicle starts at $x_0(0) = 0$ and $x_i(0) = -i\delta$ for $1 \leq i \leq N$. Let $w_i(t)$ and $u_i(t)$ denote the disturbance and control input at the ith vehicle node. The goal is to achieve string stability in this platoon, which means that small disturbances at the beginning of a chain of vehicles does not get amplified as one progresses down the chain.

The dynamics for the i-th vehicle are assumed to be linear time invariant with a scalar transfer function $H(s) \in \mathbb{C}$ and scalar input $u_i(t) \in \mathbb{R}$. The vehicle dynamics are given in the Laplace domain as

$$X_i(s) = H(s)(U_i(s) + W_i(s)) + \frac{x_i(0)}{s} \text{ for } 1 \leq i \leq N, \tag{4.1}$$

where $x_i(0)$ is the initial state of the ith system and $W_i(s)$ is the input disturbance at the ith node

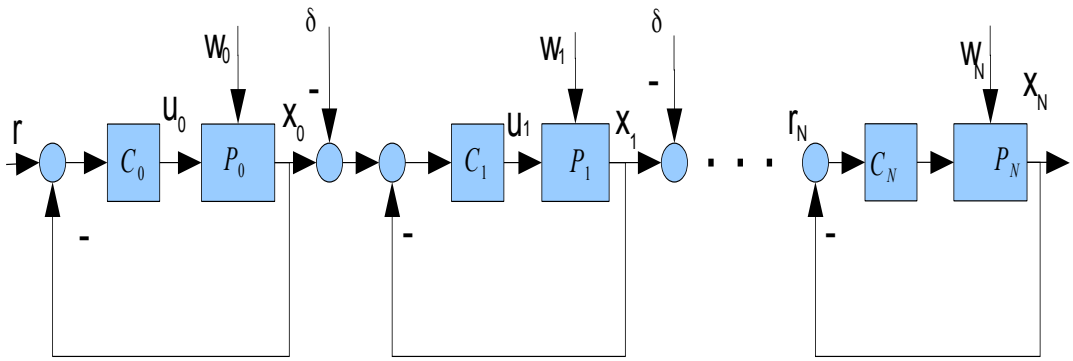


Figure 4.1: Automated Highway System platoon.

in Laplace domain.

The spacing error is defined as $e_i(t) = x_{i-1}(t) - x_i(t) - \delta$ where δ is the constant desired vehicle spacing. Let $E_i(s) = \mathcal{L}(e_i(t))$ represent the Laplace transform of the spacing error $e_i(t)$. Then, we have

$$E_i(s) = X_{i-1}(s) - X_i(s) - \frac{\delta}{s}, \quad 1 \leq i \leq N. \tag{4.2}$$

4.2.2 Network Model

In the vehicle platoon shown in Fig. 4.1, we consider three types of graph components. Assume that each follower measures the distance to its predecessor, $x_{i-1} - x_i$, such that $e_i = x_{i-1} - x_i - \delta$ is known only to the ith vehicle. Sensing graph $\mathbf{G}_s(V, E_s)$ is a fixed directed graph, whose edges are the collection of node pairs in which one measures the distance to another (its predecessor). With

this measurement, each node transmits its spacing error e_i to other followers over a communication graph, denoted by $\mathbf{G}_c(V_f, E_c)$, which could be time-varying due to fading and other environmental effects on communication channel and V_f is a set of follower nodes. Leader adjacency graph $\mathbf{G}_l(V, E_l)$ denotes each node's connection with the leader by receiving its broadcast or have a direct link with the leader, where V is a set of vertices including the leader and all followers.

By combining the three graph components, information graph $\mathbf{G}_i(V, E)$ is formed as shown, for example, in Fig. 4.2, where $E = E_l \bigcup E_s \bigcup E_c$. It is necessary that the information graph is *connected from leader node* (node 0) such that the leader information could be transmitted to each follower in the string. Since each follower measures the distance to its predecessor, the sensing graph is fixed and connected from leader node. Hence, the information graph is connected from leader node even if the communication graph is time-varying with different topologies. Let us define the Laplacian matrix of communication graph G_c as $\mathtt{L} = [l_{ij}] \in \mathbf{R}^{N\times N}$ and the leader adjacency matrix $\mathtt{C}$ as a lower triangular matrix with $c_{ij} = d_i$ for $1 \leq j \leq i$, where d_i is equal to 1 whenever the agent i is a neighbor of the leader and 0 otherwise.

4.3 Leader-Following Consensus Over a Communication Network

Consensus algorithms have been considered as a basic approach to achieve cooperative control and formation control in robotics and sensor networks [147, 148]. In a leader-following consensus problem, there is a leader that specifies the objective for the whole group. The agreement of a group of agents on their common states is to be achieved via local interaction [148, 149]. The leader is a special agent whose state is normally independent of the others and is to be followed by the others. However, since not all followers may have access to information of the leader's state, the followers may exchange their information by forming a communication network. The leader's information could be passed onto followers if there exists a directed path from the leader to these

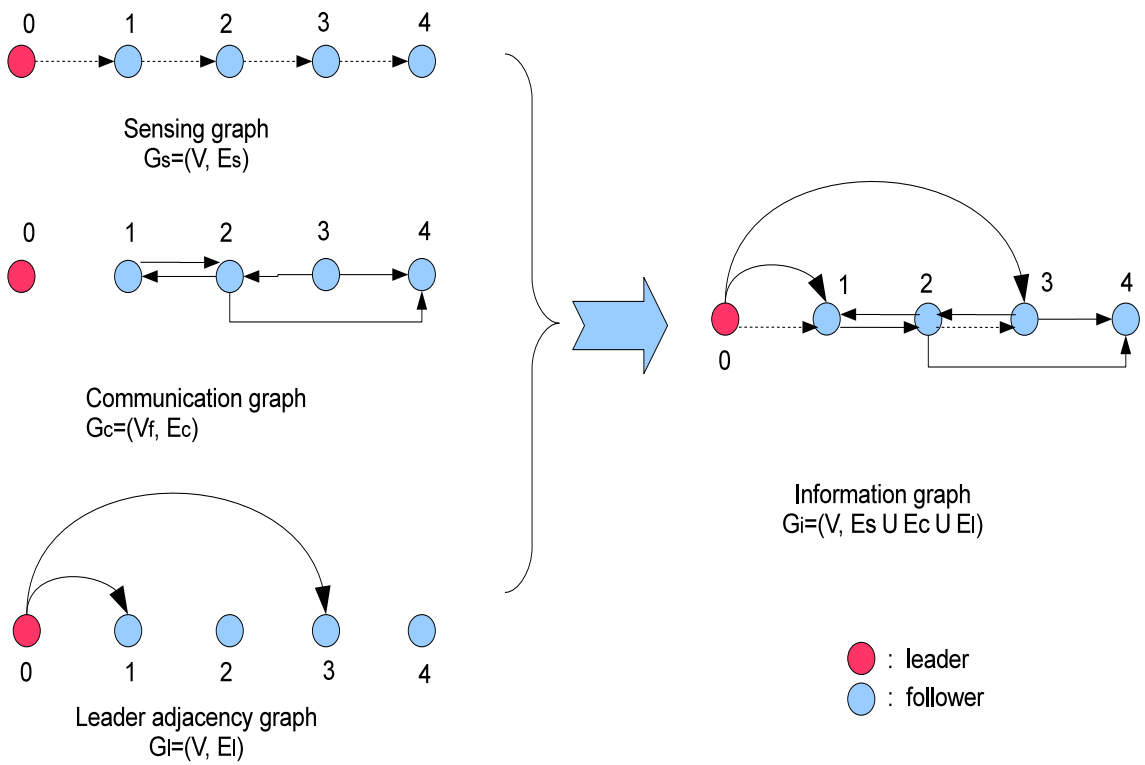


Figure 4.2: Information graph in a platoon of five vehicles.

followers. The goal of leader-following consensus problem is to reach consensus at which the error between each follower and the leader converges to zero, or equivalently $\lim_{t\to\infty} |e_i(t)| = 0$ for $\forall i \in \{1, \cdots, N\}$. If there is a directed spanning tree in the communication graph, it is known that leader-following consensus tracking algorithm will make each follower's state converge to that of the leader as time goes on [148].

In vehicular technology context, previous research has considered different control schemes to achieve string stability, such as predecessor and leader following approach, bidirectional control and time-varying headway approach [86, 146, 150]. However, there is only a little previous work

on the application of communication techniques for node communications and its effect on the string stability itself. In the following, we discuss the application of a leader-following consensus algorithm with inter-vehicle communication in achieving string stability.

4.3.1 Sufficient Conditions for Leader-following Consensus

In this section, we propose a leader-following consensus algorithm for vehicle platoon control with network model in Section 4.2.2 and the scalar system model in (4.1). In our proposed algorithm, each agent i receives information $E_j(s)$ from its neighbors $j \in \mathcal{N}_i$ in communication graph G_c and information $X_0(s)$ from the leader if the agent i is a neighbor of the leader. Here the neighborhood $\mathcal{N}_i$ of node i is the set of vertices that send information to node i.

Each agent i calculates the difference between the spacing errors of itself and jth neighbor for $j \in \mathcal{N}_i$ and its distance from the leader and uses them to form a control input $U_i(s)$ as follows:

$$U_i(s) = K \sum_{j \in \mathcal{N}_i} l_{ij}\big(E_i(s) - E_j(s)\big) - K d_i\big(X_i(s) - X_0(s) - i\frac{\delta}{s}\big), \tag{4.3}$$

where K is a control parameter. In (4.3), the first term describes the relative position error between itself and its neighbors and the second term describes the ith node's distance from the leader when it is directly connected with the leader. If node i has a direct link with the leader, it will incorporate the leader information in the second term. Otherwise, in the first term, it can extract the leader information from the spacing errors received from its neighbors which are directly connected with the leader. Agent i needs the information $X_j(s)$ from its neighbors and leader's information $X_0(s)$ if possible. For agent i to acquire this information, we make the following two assumptions: First, each vehicle is capable of communication and a communication network among followers exists. Second, the information graph is connected from leader node such that the leader has directed paths to all followers through communication in the network [1].

[1] In the ideal case, if the leader broadcasts its position $X_0(s)$ to every follower, leader-follower consensus could easily be achieved. However, it comes at the cost of increased communication requirements as the platoon length increases. In our case, when the position of the leader $X_0(s)$ is only available to some agents, other agents can attain this information if the graph is connected from leader node.

There are various detailed models for vehicle dynamics and most are complex higher order system models [151], for which analytical results are difficult to obtain. Since in this paper we are interested in obtaining analytical conditions for string stability, we consider a simple point mass model, formulate it as a first order system and consider the proper transfer function $H(s) = \frac{B}{s-A}$ in (4.1) for the following analysis[2]. Substituting (4.3) into (4.1), we have the ith vehicle's system dynamics as

$$X_i(s) = H(s)K\sum_{j\in\mathcal{N}_i} l_{ij}(E_i(s) - E_j(s)) \tag{4.4}$$
$$- H(s)Kd_i(X_i(s) - X_0(s) - i\frac{\delta}{s}) + H(s)W_i(s) + \frac{x_i(0)}{s},$$

where $H(s) = \frac{B}{s-A}$. The error dynamics for $i = 1$ could be written as (using (4.2)),

$$E_1(s) = X_0(s) - H(s)K\sum_{j\in\mathcal{N}_1} l_{1j}(E_1(s) - E_j(s)) - H(s)(Kd_1E_1(s) + W_1(s)).$$

and for $i > 1$,

$$E_i(s) = H(s)K\big[\sum_{l\in\mathcal{N}_{i-1}} l_{i-1,l}(E_{i-1}(s) - E_l(s)) - \sum_{j\in\mathcal{N}_i} l_{ij}(E_i(s) - E_j(s))$$
$$+ d_{i-1}\sum_{k=1}^{i-1} E_k(s) - d_i\sum_{q=1}^{i} E_q(s)\big] + H(s)(W_{i-1}(s) - W_i(s)). \tag{4.5}$$

We may obtain a matrix form of error dynamics from (4.5):

$$E(s) = -H(s)K\mathbb{V}E(s) + H(s)MW(s) + \Phi_1X_0(s), \tag{4.6}$$

where $E(s) = [E_1(s)\ \cdots\ E_N(s)]^T$, $W(s) = [W_1(s)\ \cdots\ W_N(s)]^T$, $\Phi_1 = [1\ 0\cdots0]^T \in \mathbf{R}^N$, $M = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix}$, $\mathbb{V} = \mathtt{L} - \tilde{\mathtt{L}} + \mathtt{C} - \tilde{\mathtt{C}}$, $\tilde{\mathtt{L}}$ is a shifted version of matrix $\mathtt{L}$ such that

[2]The chosen transfer function is a strictly proper transfer function of a linear time-invariant system, which is obtained by feedback linearization of an abstraction of vehicle model [146].

$\tilde{l}_{1,j} = 0$ and $\tilde{l}_{i,j} = l_{i-1,j}$ for $2 \leq i \leq N$ and $1 \leq j \leq N$, $\tilde{\mathbb{C}}$ is a shifted version of matrix $\mathbb{C}$ such that $\tilde{c}_{1,j} = 0$ and $\tilde{c}_{i,j} = c_{i-1,j}$ for $2 \leq i \leq N$ and $1 \leq j \leq N$. From (4.6), we have

$$E(s) = G_{de}(s)W(s) + G_{xe}(s)X_0(s), \tag{4.7}$$

where $G_{de}(s) = [1 + H(s)K\mathbb{V}]^{-1}H(s)M$ represents the transfer function matrix from disturbance $W(s)$ to error $E(s)$ and $G_{xe}(s) = [1 + H(s)K\mathbb{V}]^{-1}\Phi_1$ represents the transfer function matrix from leader position $X_0(s)$ to error $E(s)$.

We are interested in analyzing the error prorogation as the length N of the platoon increases. As mentioned earlier, string stability requires that spacing errors do not amplify as they propagate upstream from one vehicle to another vehicle [146, 150]. Here we modify the definition of string stability from [150] and restate it as string stability with respect to input disturbance $w(t)$ as follows:

Definition 4.3.1. ***(String stability):*** *Consider a string of N dynamic systems described in (4.1) with error dynamics given by (4.2). Let $W_i(s)$ be the input disturbance at the ith vehicle. The string is stable with respect to input disturbance if given any $\epsilon > 0$ there exists a $\delta > 0$, such that $||W_i(s)||_\infty < \delta \quad \Rightarrow ||E_i(s)||_\infty < \epsilon \quad$ for $\quad i \in \{1, \cdots, n\}$, where $||W_i(s)||_\infty = \sup_{\omega \in \mathbb{R}} \sigma_{\max}(W_i(j\omega))$ is the H_∞ norm of transfer function matrix $W_i(s)$ and $\sigma_{\max}$ is the maximum singular value [6].*

The above definition states that if the input disturbance of each vehicle is uniformly bounded, the error dynamics of each follower vehicle must also be uniformly bounded. From (4.7), $X_0(s)$ is constant and $W(s)$ is uniformly bounded. To achieve bounded error $E(s)$, we need the largest disturbance to error transfer function which is the largest element of transfer function matrix $G_{de}(s)$ to be uniformly bounded. *In other words, if $||G_{de}||_\infty = \sup_\omega \sigma_{\max}\big(G_{de}(j\omega)\big)$ is uniformly bounded, then the string is stable. It means that $\sigma_{\max}\big(G_{de}(j\omega)\big)$ must be uniformly bounded over all frequencies ω.*

In the following, we invoke a lemma on H_∞ norm bound from [152] and then obtain our main theorem on the sufficient conditions for string stability by bounding the disturbance to error transfer function matrix G_{de}:

Lemma 4.3.2. *For the following continuous time LTI system*

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= A\boldsymbol{x}(t) + B\boldsymbol{u}(t),\\ \boldsymbol{y}(t) &= C\boldsymbol{x}(t),\end{aligned} \tag{4.8}$$

if the solution of the nonstandard Riccati equation

$$0 = A^T P + PA + C^T C + PBB^T P \tag{4.9}$$

is positive definite and the pair (A, C) is observable, then the following H_∞ norm bound is satisfied

$$||C(SI - A)^{-1}B||_\infty \leq 1. \tag{4.10}$$

Proof. See the proof in [152]. □

Lemma 4.3.2 shows that the H_∞ norm of system transfer matrix is bounded if the nonstandard Riccati equation has positive definite solutions. Meanwhile, the H_∞ norm bound is arbitrary. The next lemma shows the properties of matrix M in (4.6)

Lemma 4.3.3. *The maximal singular values of matrix M defined as* $M = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix}$ *is less than 2. i.e. $\sigma_{\max}(M) \leq 2$. The eigenvalues of M are equal to -1. i.e. $\lambda_i(M) = -1$ for all i.*

Proof. See Appendix A. □

By adopting Lemma 4.3.2 and 4.3.3, we will obtain the following theorem.

Theorem 4.3.4. ***(Suffcient conditions for string stability):*** *Consider a platoon of vehicles as depicted in Fig. 1 with vehicle dynamics given by (4.1) with $H(s) = \frac{B}{S-A}$. Each vehicle is to use the leader-following consensus algorithm given in (4.3). The transfer function matrix $G_{de}(s)$ from*

disturbance $W(s)$ to error $E(s)$ is given in (4.7), where matrices $\mathtt{L}$ and $\mathtt{C}$ are the communication graph Laplacian matrix and leader adjacency matrix, respectively. If the solution of the nonstandard Riccati equation

$$0 = (A - BK\mathbb{V})^T P + P(A - BK\mathbb{V}) + I + PBB^T P \tag{4.11}$$

is positive definite, where $\mathbb{V} = \mathtt{L} - \tilde{\mathtt{L}} + \mathtt{C} - \tilde{\mathtt{C}}$, then $||G_{de}||_\infty < 2$ for all $N \in \{1, 2, \cdots\}$.

Proof. Consider the sufficient conditions for $||G_{de}(s)||_\infty \leq 1$[3], where $G_{de}(s) = [I+H(s)K\mathbb{V}]^{-1}H(s)M$. From the sup-multiplicative property of H_∞ norm [156], we have $||G_{de}(s)||_\infty \leq ||T(s)||_\infty ||M||_\infty$, where $T(s) = [I + H(s)K\mathbb{V}]^{-1}H(s)$ and, as assumed earlier, $H(s) = \frac{B}{s-A}$. From Lemma 4.3.3, we have that $\sigma_{\max}(M) \leq 2$, which yields $||M||_\infty \leq 2$ for M is a constant matrix. With this, we have

$$||G_{de}(s)||_\infty \leq ||T(s)||_\infty ||M||_\infty \leq 2||T(s)||_\infty \tag{4.12}$$

Recall that $||T(s)||_\infty = |||[I + H(s)K\mathbb{V}]^{-1}H(s)||_\infty$ and substitute $H(s) = \frac{B}{s-A}$ into $T(s)$. Then we have

$$||T(s)||_\infty = |||[sI - (A - BK\mathbb{V})]^{-1}B||_\infty. \tag{4.13}$$

Let us formulate the following LTI system with $T(s)$ as its transfer function

$$\begin{aligned} \dot{\mathbf{x}}(t) &= (A - BK\mathbb{V})\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{x}(t). \end{aligned} \tag{4.14}$$

Since output matrix $C = I$, it is clear that the pair (C, A) is observable. By using Lemma 1, if the solution of (4.11) is positive definite, then $||T(s)||_\infty \leq 1$. Later, by substituting this into (4.12), we will have $||G_{de}(s)||_\infty \leq 2$. □

[3]Because we want to achieve an analytical condition for simulation purpose, we choose a specific bound $||G_{de}|| \leq 2$ which could be easily extended to a general case $||G_{de}|| \leq \epsilon$, for given $\epsilon > 0$.

Theorem 4.3.4 shows that if the graph Laplacian matrix $\mathtt{L}$, leader-adjacency matrix $\mathtt{C}$ and control parameter K satisfy (4.11), then the spacing error does not amplify as it propagates in the string regardless of the length of the platoon N. The result of Theorem 4.3.4 implies that by applying the leader-consensus following algorithm, the vehicle platoon is robust to disturbances in the string and string stability is guaranteed regardless of the size N of the platoon. The sufficient conditions given in (4.11) can thus provide guidelines for adjusting the control parameter K in each vehicle and the topology design of the communication network in the platoon so as to ensure string stability.

4.3.2 Comparison with Previous Approaches

There are several control approaches previously proposed in literature to achieve string stability and vehicle platoon control. However, they mainly focus on control performance and stability analysis without mentioning the underlying communication graph. By proposing the leader-following consensus algorithm with a general communication graph, we could consider the previous control approaches as special cases of our proposed framework with specific types of assumed communication graphs.

Predecessor Following Control

Predecessor following control is a linear control law based on relative spacing error with respect to the predecessor [8]. Each vehicle only receives information from its predecessor without direct access to the leader. In [8], a linear, single-input-single-output system model with first order actuator dynamics is considered. The control method of predecessor following control is

$$U_i(s) = K_f(s)E_i(s), \tag{4.15}$$

with the scalar spacing error dynamics as

$$E_1(s) = S(s)X_0(s) = \frac{1}{1 + H(s)K_f(s)}X_0(s)$$
$$E_i(s) = T(s)E_{i-1}(s) = \frac{H(s)K_f(s)}{1 + H(s)K_f(s)}E_{i-1}(s) \tag{4.16}$$

where the sensitivity function $S(s)$ is the transfer function from $X_0(s)$ to $E_1(s)$ and complementary sensitivity function $T(s)$ is the transfer function from $E_{i-1}(s)$ to $E_i(s)$.

In the predecessor following method, the leader information is not broadcasted and each follower only senses its predecessor. From Section 4.2.2, we can see that the leader adjacency graph is not connected and $\mathtt{C} = 0$. Since there is no communication between every two followers, the communication graph is not connected and communication graph Laplacian matrix $\mathtt{L} = \begin{bmatrix} 1 & 0 & \cdots & \\ 0 & 1 & \cdots & \cdots \\ & \ddots & \ddots & 0 \\ \cdots & \cdots & 0 & 1 \end{bmatrix}$. By using (4.6), we can obtain that $\mathtt{V} = \begin{bmatrix} 1 & 0 & \cdots & \\ -1 & 1 & \cdots & \cdots \\ & \ddots & \ddots & 0 \\ \cdots & \cdots & -1 & 1 \end{bmatrix}$.

Substitute matrix $\mathtt{V}$ in (4.11) with other parameters shown in Section 4.4 and solve the algebraic riccati equation (ARE). The Hamiltonian matrix $P_h = \begin{bmatrix} A - BK\mathtt{V} & BB^T \\ -I & -(A - BK\mathtt{V})^T \end{bmatrix}$ has eigenvalues on the imaginary axis. Hence, the ARE does not have stabilizing solutions from Lemma A.2.2 in [153]. From the result of Theorem 4.3.4, there does not exist bounded H_∞ norm for predecessor following approach. Hence, the predecessor following approach could not guarantee string stability.

Predecessor and Leader Following Control

In predecessor and leader following control [8], each vehicle receives information from its predecessor and the leader will broadcast its location to each follower. With this information, each vehicle applies a linear control law based on relative spacing errors with respect to the predecessor

and the leader, respectively:

$$U_i(s) = K_p(s)E_i(s) + K_l(s)\big(X_0(s) - X_i(s) - \frac{i\delta}{s}\big). \tag{4.17}$$

The scalar spacing error dynamics is written as

$$\begin{aligned} E_1(s) &= S_{lp}(s)X_0(s) = \frac{1}{1 + H(s)(K_p(s) + K_l(s))}X_0(s), \\ E_i(s) &= T_{lp}(s)E_{i-1}(s) = \frac{H(s)K_p(s)}{1 + H(s)(K_p(s) + K_l(s))}E_{i-1}(s). \end{aligned} \tag{4.18}$$

From Section 4.2.2, since each vehicle receives information from the leader, we have $d_i = 1$ for $i = 1, \cdots, N$ and $\mathtt{C} - \tilde{\mathtt{C}} = I$. Because of no communication among followers, the communication graph is not connected and communication graph Laplacian matrix $\mathtt{L} = \begin{bmatrix} 1 & 0 & \cdots & \\ 0 & 1 & \cdots & \cdots \\ & \ddots & \ddots & 0 \\ \cdots & \cdots & 0 & 1 \end{bmatrix}$.

Then, we obtain matrix $\mathtt{V} = \begin{bmatrix} 2 & 0 & \cdots & \\ -1 & 2 & \cdots & \cdots \\ & \ddots & \ddots & \\ \cdots & \cdots & -1 & 2 \end{bmatrix}$

Substituting matrix $\mathtt{V}$ in (4.11) with other parameters shown in Section 4.4 and solving this algebraic riccati equation (ARE), we can find out that predecessor and leader following approach has unique positive stabilizing solutions. From the result of Theorem 4.3.4, there exists bounded H_∞ norm for predecessor and leader following approach, which is consistent with the simulation results in Section 4.4.

Bidirectional Control with Leader Information

In bidirectional control with leader information [154, 155], platoon controller uses relative spacing errors with respect to adjacent vehicles (predecessor and follower) and the leader. Consider the

bidirectional control strategy with leader information as follows: for $i = 1, \cdots, N-1$,

$$U_i(s) = K_p(s)(E_i(s) - E_{i+1}(s)) + K_l(s)(X_0(s) - X_i(s) - \frac{i\delta}{s}).$$

For $i = N$,

$$U_N(s) = K_p(s)E_N(s) + K_l(s)(X_0(s) - X_N(s) - \frac{N\delta}{s}).$$

Here we assume that the spacing error of the leader is $E_0(s) = 0$ and the leader has the same controller as other sub-systems. Then, we have the scalar spacing error as For $1 \leq i \leq N-1$

$$\begin{aligned} E_i(s) &= H(s)K_p(s)[E_{i-1}(s) + E_{i+1}(s) - 2E_i(s)] \\ &\quad - H(s)K_l(s)E_i(s) + H(s)(W_{i-1}(s) - W_i(s)). \end{aligned}$$

For $i = N$,

$$\begin{aligned} E_N(s) &= H(s)K_p(s)[E_{N-1}(s) - 2E_N(s)] \\ &\quad - H(s)K_l(s)E_N(s) + H(s)(W_{N-1}(s) - W_N(s)). \end{aligned}$$

In vector form, we have

$$\begin{aligned} \overline{E}(s) &= G_{de}(s)\overline{W}(s) + G_{xe}(s)\overline{X}_0(s) \\ &= [1 + H(s)K_p(s)\mathbb{V}]^{-1}H(s)M\overline{W}(s) + [1 + H(s)K_p(s)\mathbb{V}]^{-1}\Phi_1\overline{X}_0(s), \end{aligned} \tag{4.19}$$

where

$$\mathbb{V} = \begin{bmatrix} 2 & -1 & & \\ -1 & 3 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 3 \end{bmatrix}, M = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix}.$$

Through comparison with leader-following consensus algorithm with a general communication graph, we have $\mathbb{C} - \tilde{\mathbb{C}} = I$ and $\mathbb{L} = \begin{bmatrix} 1 & -1 & \cdots & \\ & 1 & -1 & \cdots \\ & & \ddots & \ddots \\ \cdots & & & 1 \end{bmatrix}$, since the ith vehicle communicates

with the $i+1$th vehicle and the leader information is broadcasted to each follower. Then, we have the same $\mathtt{V}$ as in (4.19).

Substituting matrix $\mathtt{V}$ in (4.11) with other parameters shown in Section 4.4 and solving this algebraic riccati equation (ARE), we can find out that bidirectional control with leader information approach has unique positive stabilizing solutions. From the result of Theorem 4.3.4, there exists bounded H_∞ norm for bidirectional control with leader information approach, which is consistent with the simulation results in Section 4.4.

The results obtained from above analysis is consistent with those in [8, 154, 155] and we are able to fit these previously control schemes into different specific communication topologies in our scenario. In the next section, we will study the performance of those approaches and compare them with our proposed leader-follower consensus algorithm.

4.4 Simluation Results

In this section, we study the performance of the proposed leader-following consensus algorithm under different communication graphs. We will also compare the performance with previously proposed approaches described in Section 4.1: predecessor following, predecessor and leader following, and bidirectional control with leader information.

In the first simulation, we investigate the performance of leader-following consensus algorithm with different communication graphs. We consider a scalar vehicle model in (4.1). We consider that each follower is connected to leader and the communication graph among followers is a random graph with time-invariant topology $G(N,p)$, where $0 \leq p \leq 1$ is the probability of having a link between any two nodes. The assumed parameters are: $A=1$, $B=2$, $K=1.5$ and $N=1,2,5,10,20$. Figures 4.3, 4.4 and 4.5 show the performance of the leader-following consensus algorithm with random graphs for $p=0.1, 0.5$ and 1, respectively. Here x-axis represents the frequency ω and y-axis represents the largest singular value $\sigma_{\max}\big(G_{de}(j\omega)\big)$ of disturbance to

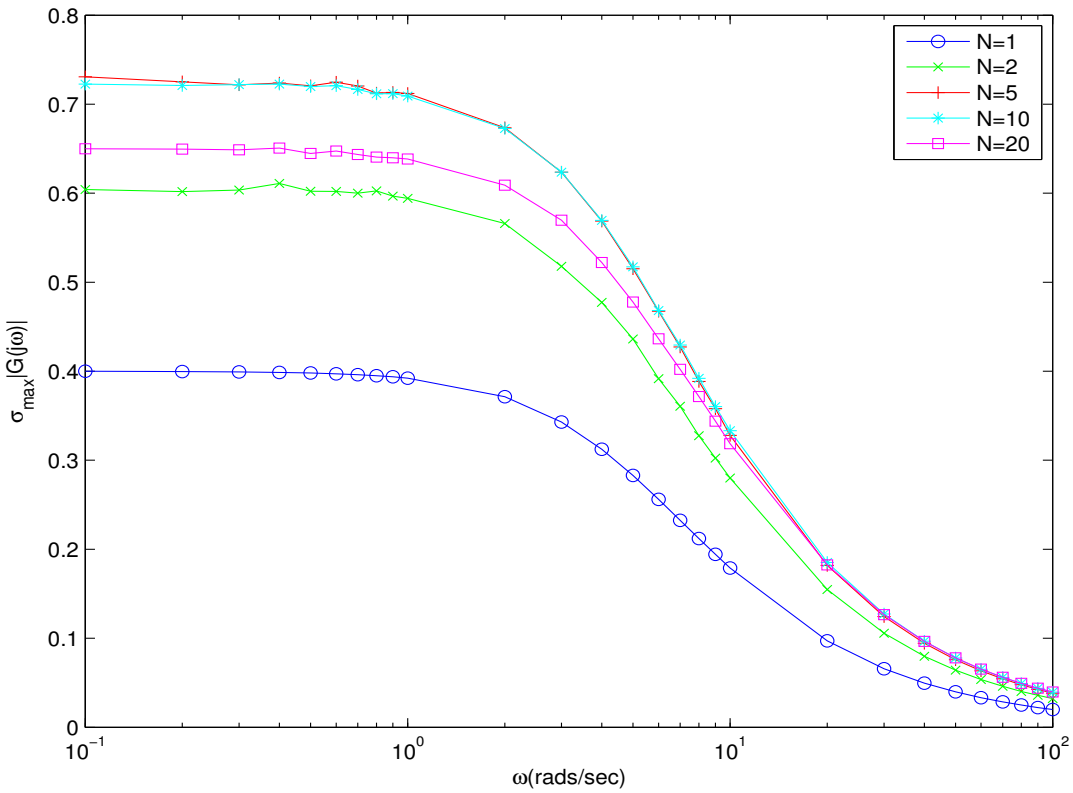


Figure 4.3: The performance of leader-following consensus algorithm with random communication graph among followers – $G(N,p)$ ($p = 0.1$).

error transition matrix $G_{de}(j\omega)$ at frequency ω. As seen from Fig. 4.3, since each follower is connected to leader, with proper parameters, even if the communication graph is less connected, the string stability is achieved (disturbance to error gain $\sigma_{\max}\big(G_{de}(j\omega)\big)$ is bounded at each ω in frequency domain.). From Figs. 4.3, 4.4 and 4.5, it can be seen that for the same string length N, the disturbance to error gain $\sigma_{\max}\big(G_{de}(j\omega)\big)$ decreases as the graph connectivity increases. Compare the peak values of curves in Fig. 4.4 and Fig. 4.5. In Fig. 4.4, the peak value of each curve $||G_{de}||_\infty = \sup_\omega \sigma_{\max}\big(G_{de}(j\omega)\big)$ first increases then decreases as N increases. However, in Fig. 4.5, the peak value decreases as N increases. The reason is due to the combined effect of platoon size N and link probability p on error propagation in the platoon. If the platoon size N increases, then the error propagates longer, which tends to increase the maximum spacing error in the string.

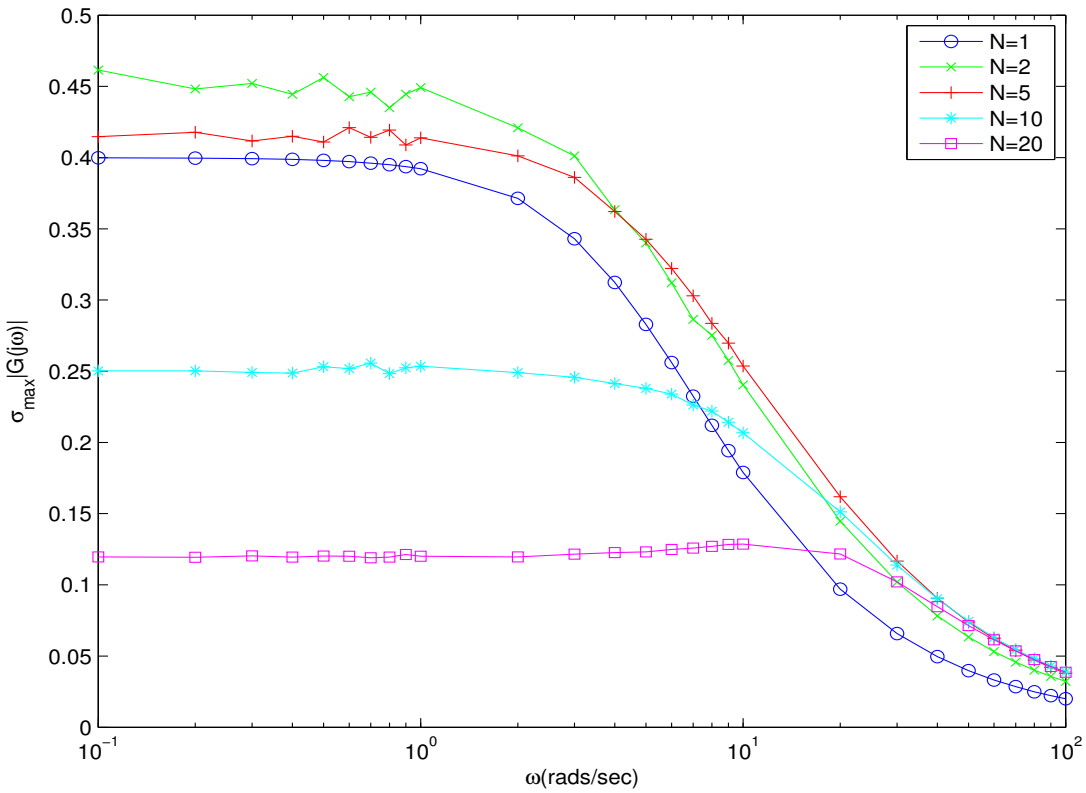


Figure 4.4: The performance of leader-following consensus algorithm with random communication graph among followers – $G(N, p)$ ($p = 0.5$).

However, as link probability p increases, the graph becomes more connected, which mitigates the effect of error propagation and leads to lower spacing error. Therefore, we can see the from Fig. 4.4 that with lower link probability, the string length plays a dominating effect and it leads to higher spacing error (higher value of $||G_{de}||_\infty$). In Fig. 4.5 with higher link probability, the link probability plays a dominating effect and it leads to lower spacing error (lower value of $||G_{de}||_\infty$).

In the second simulation, we compare the performance of our approach and that of previously proposed approaches [146, 150]: predecessor following, predecessor and leader following, and bidirectional control with leader information. We consider the same scalar vehicle model as before: $A = 1$, $B = 2$ with $K_f = 3$, $Kp = K_l = K = 1.5$, $p = 0.5$ and $N = 1, 2, 5, 10, 20$, where K_f,

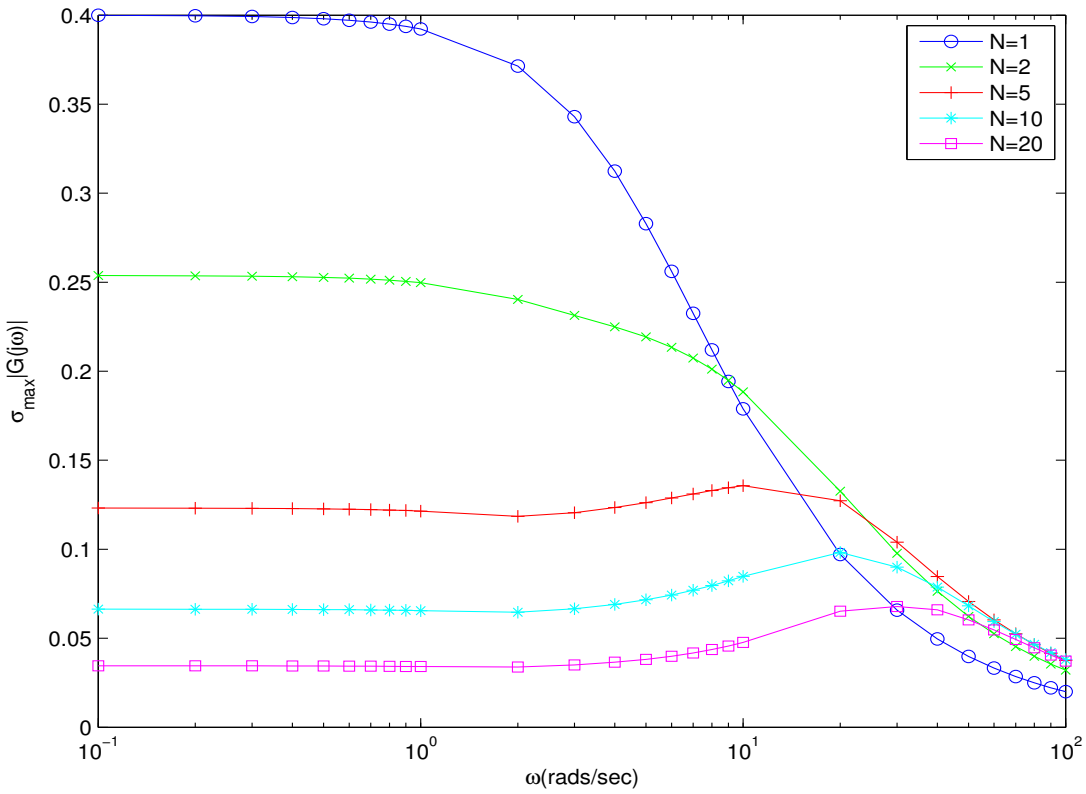


Figure 4.5: The performance of leader-following consensus algorithm with random communication graph among followers – $G(N, p)$ ($p = 1$).

K_p, K_l are the control parameters associated with predecessor following, leader and predecessor following and bidirectional control with leader information approaches, respectively [146]. All three approaches are with fixed communication graphs according to their control schemes. Figures 4.6, 4.7 and 4.8 show the performance of predecessor following, predecessor and leader following and bidirectional control with leader information, respectively. In Fig. 4.8, the disturbance to error gain is relatively independent of platoon size N. In Fig. 4.7, the disturbance to error gain increases as N grows but it is still bounded when N is large, which is consistent with the results reported in [146]. Here the peak value of curves at frequency $\omega \neq 0$ is due to the control scheme it applies [146]. We see from Figs. 4.7 and 4.8 that bidirectional control method performs better than predecessor and leading following method because the former assumes stronger connectivity

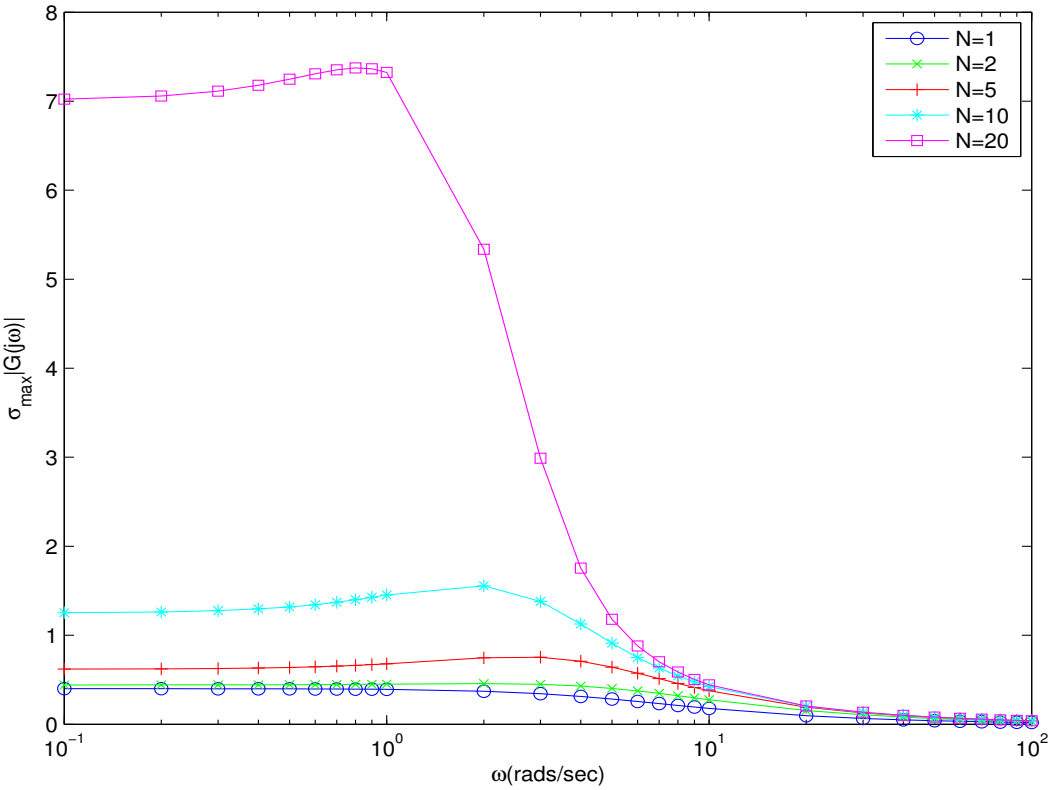


Figure 4.6: The performance of predecessor following approach.

in the communication graph. Indeed, in bidirectional control method, each agent i communicates with its preceding and following vehicles, while in predecessor and leader following method it only receives information from its previous vehicle. From Fig. 4.6, it can be seen that simple predecessor following method may not guarantee string stability as error propagation may grow unbounded with large platoon size.

By comparing Figs. 4.6-4.8 with Fig. 4.4, we can see that compared to previously proposed methods, our approach achieves a similar disturbance to error gain for smaller platoon size (N=5) but achieves a smaller disturbance to error gain for large platoon size (N=20). Thus, for example, with average link connectivity $p = 0.5$, our approach performs better than previously proposed approaches for large platoon size. Moreover, previously proposed methods require that each vehicle

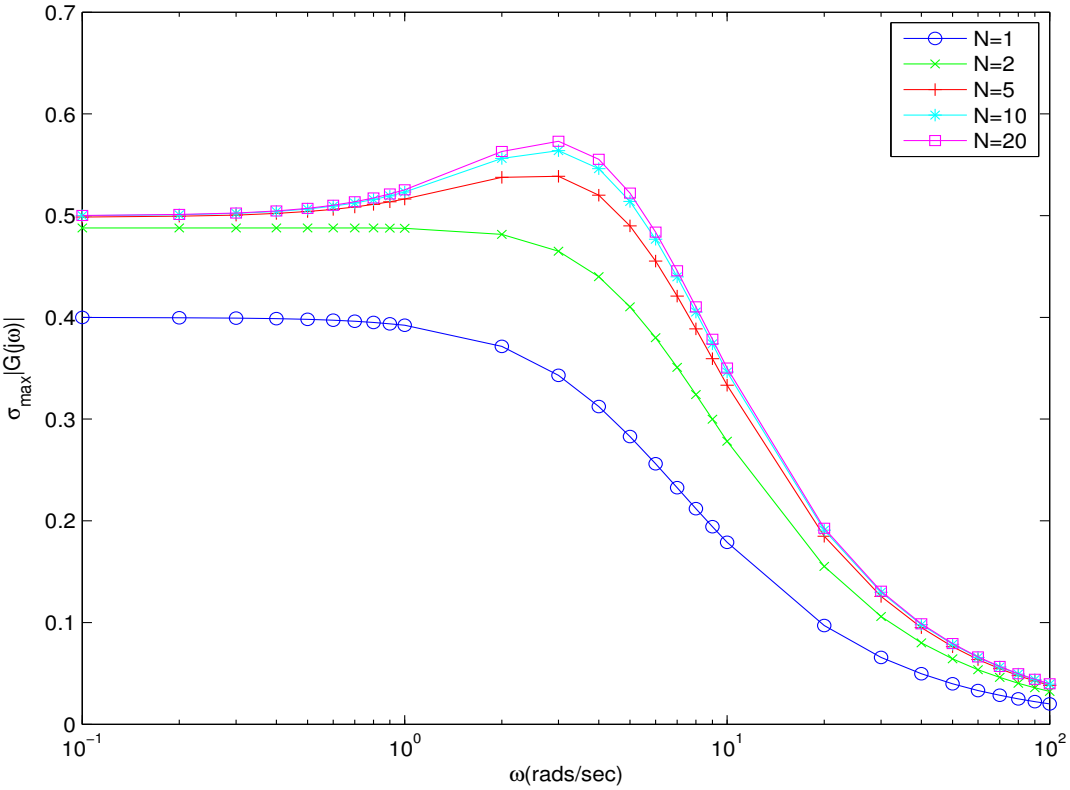


Figure 4.7: The performance of predecessor and leader following approach.

must know its position in the string in order to apply associated control schemes. The controller has different forms for the first vehicle, the last one and the rest in between. In our proposed algorithm, no information about the number of cars in the string is needed because each vehicle only communicates with its neighbors and processes information locally. This allows much simpler coordination requirements compared to previously proposed methods and makes it more scalable and distributed. Our proposed algorithm applies the same controller form to each vehicle, which also leads to simplicity in implementation of on-vehicle controller.

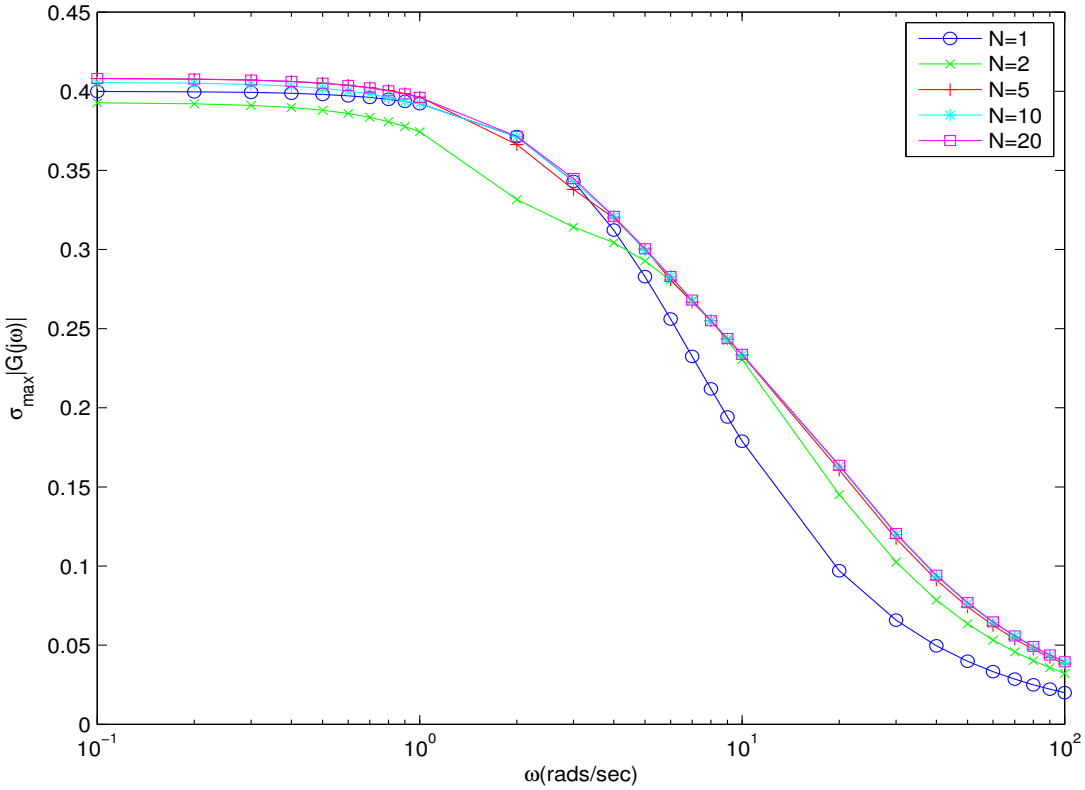


Figure 4.8: The performance of bidirectional control with leader information approach.

4.5 Conclusions

In this chapter we introduced a leader-following consensus algorithm with a communication network to achieve string stability in vehicle platoons. We showed the sufficient conditions on the connectivity of communication graph and the control parameters in each vehicle for string stability by using the leader-following consensus algorithm. From simulation results, we investigated the effects of communication topology on string stability in terms of disturbance propagation. We demonstrated the advantage of our proposed approach in term of distributed processing, scalability and simpler coordination requirements by comparing its performance with the previously proposed control methods. Future research may further analyze tradeoff between the graph connectivity and

control cost and apply these results into controller design and communication topology design.

In the next chapter, we consider the problem of distributed tracking in wireless sensor networks (WSN) through cooperative control. We formulates the problem of the distributed tracking in a network of sensors with a time-varying network topology, incomplete node data and noisy communication links. By using graph theory and cooperative control, we propose a distributed tracking with consensus algorithm to improve the cooperative tracking performance of the network and we also prove the convergence of the proposed algorithm and analyze its steady state behavior.

Chapter 5

Distributed Tracking and Consensus over Networked Multiagent Systems

5.1 Introduction

In this chapter, we consider distributed tracking with consensus over networked multi-agent systems which were proposed in [119]. Distributed tracking with consensus refers to the problem that a group of nodes need to achieve an agreement over the state of a dynamical system by exchanging tracking estimates over a network. Information exchange among nodes may improve the quality of local estimates and consensus estimates may help avoid conflicting and inefficient distributed decisions. Other applications of this problem include flocking and formation control, real-time monitoring, target tracking and global positioning system (GPS) [119, 120]. In the rest of this chapter, we provide proof to the results in [121, 122] and analyze the convergence of the proposed algorithm and the steady state behavior. The outline of this chapter is as follows. Section 5.2 introduces our assumed system and network model and the proposed distributed tracking with consensus algorithm. In Section 5.3 conditions for achieving distributed consensus are discussed and the rate of convergence is quantified. The steady-state performance of the proposed distributed

tracking with consensus algorithm is also analyzed in Section 5.3 . Section 5.4 provides detailed simulation results and performance comparison of the proposed distributed tracking with consensus algorithm and that of distributed local Kalman filtering with centralized fusion and centralized Kalman filter. Finally, conclusions are given in Section 5.5 .

5.2 Tracking over Noisy Time-varying Graphs with Incomplete Data

5.2.1 Problem Formulation

Consider an N-node sensor network with a *connectivity graph* $G(j) = \big(V, E(j)\big)$ at time j. Assume that the graph $G(j)$ is undirected, but time varying due to nodes moving out of communication ranges of each other or needing to cease transmissions to save battery power. The objective is to perform distributed tracking of a target and exchange tracking estimates over noisy communication links and try to reach consensus over the network. The tracking updates are performed at k instances, where k denotes the tracking time step $(k = 0, 1, \cdots)$. Consensus updates are performed between every two tracking updates, where $0 \leq j < J$ denotes the consensus iteration number and J is the number of consensus iterations per tracking update (assumed to be fixed). The dynamics of the target evolves according to

$$\mathbf{x}(k+1) = F\mathbf{x}(k) + \mathbf{w}(k); \quad \mathbf{x}(0) \sim \mathcal{N}\big(\overline{\mathbf{x}}(0), P_0\big). \tag{5.1}$$

The sensing model of the n-th sensor is

$$\mathbf{y}_n(k) = H_n\mathbf{x}(k) + \mathbf{v}_n(k), \quad \mathbf{y}_n \in \mathbb{R}^l. \tag{5.2}$$

Note that, the observation matrices H_n's can be different for each node. Both $\mathbf{w}(k)$ and $\mathbf{v}_n(k)$ are assumed to be zero-mean white Gaussian noise (WGN) and $\mathbf{x}(0) \in \mathbb{R}^M$ is the initial state of the

target. The second-order statistics of the process and measurement noise are given by

$$\mathbb{E}\big[\mathbf{w}(k)\mathbf{w}(k')^T\big] = Q\delta_{kk'}, \qquad \mathbb{E}\big[\mathbf{v}_n(k)\mathbf{v}_{n'}(k')^T\big] = R_n\delta_{kk'}\delta_{nn'},$$

where $\delta_{kk'} = 1$ if $k = k'$ and $\delta_{kk'} = 0$, otherwise. Note that the above system model is linear, while the system model assumed in [119] is highly nonlinear making it difficult to analyze to obtain theoretical performance characterization.

Figure 5.1 shows the system model of distributed tracking with consensus on a time-varying graph with incomplete data and noisy communication links. Let $\overline{\mathbf{x}}_n(k,j)$ denote the node n's updated tracking estimate at the j-th consensus iteration that follows the k-th tracking update step with $\overline{\mathbf{x}}_n(k,0) = \hat{\mathbf{x}}_n(k|k)$, where $\hat{\mathbf{x}}_n(k|k)$ is the n-th node's filtered tracking estimate in the k-th tracking update. The received data at node n from node l, for $n \neq l$, at iteration j can be written as

$$\mathbf{z}_{n,l}(k,j) = \overline{\mathbf{x}}_l(k,j) + \boldsymbol{\phi}_{n,l}(j), \text{ for } 0 \leq j < J, \tag{5.3}$$

where $\boldsymbol{\phi}_{n,l}(j)$ denotes the receiver noise at the node n in receiving the estimate of node l at iteration j and $\mathbf{z}_{n,n}(k,j) = \overline{\mathbf{x}}_n(k,j)$. Assume that $\mathbb{E}\big[\boldsymbol{\phi}_{n,l}(j)\big] = \mathbf{0}_M$, $\mathbb{E}\big[\boldsymbol{\phi}_{n,l}(j)\boldsymbol{\phi}_{n,l}^T(j)\big] = \sigma_{n,l}^2 I_M$ with $\sup_{n,l,j}\mathbb{E}\big[\|\boldsymbol{\phi}_{n,l}(j)\|^2\big] = u < \infty$.

As depicted in Fig. 5.1, at the end of the k-th tracking update, each node n which has an observation of the target will have a filtered estimate $\hat{\mathbf{x}}_n(k|k)$ with associated covariance matrix $\hat{P}_n(k|k)$. In order to improve the tracking estimate accuracy, it will exchange this filtered estimate with its neighbors over noisy communication links and try to reach consensus over the network. Note that, the goal here is to obtain a consensus tracking estimate over the local estimates at each tracking time step k, and thus, the consensus problem is essentially a problem of *consensus in estimation*.

Due to time-varying topology of the network, at any given tracking time step k not all nodes may observe the target. Thus, these nodes will not have local tracking estimates, which is denoted as *incomplete data*. In previous consensus literature [14,99,101–111], all node estimates are taken

into account in forming consensus estimates. However, the same method may not be extended to incomplete data case, since the nodes that mostly do not have observation ($\mathbf{y}_n(k) = \mathbf{v}_n(k)$) will exchange their predicted filtered estimates with others. Those predicted tracking estimates are considered as valid estimates and are taken into account to form consensus estimates, which results in inaccurate estimates and worsens the sensor network performance. By considering incomplete data here, the nodes do not have data will not communicate their invalid tracking estimates (by setting $\hat{\mathbf{x}}_n(k|k) = \mathbf{0}$ and $\hat{P}_n(k|k) = \epsilon I_M$ for some $\epsilon > 0$ instead). By introducing active node set and effective network graph, each node will notice which node has data in current consensus iteration. Only the estimates from active nodes are considered into forming consensus estimates. The estimates from non-active nodes will not be considered until it forms its updated estimate by fusing the filtered estimates from neighboring active nodes. Since the non-active nodes join the consensus process without invalid tracking estimates, faster consensus process could be achieved while the network performance is still maintained.

In the space object tracking problem treated in [119], each node observes the target and locally processes its data in *data sampling period*. After forming local estimates, each node will share its information among neighboring nodes in *information sharing period*. Here the information sharing rate is much larger compared to the data sampling rate so that each data sample node may exchange their local estimates many times in between, which may conceivably lead to better consensus estimates. The distributed tracking with consensus problem as formulated above may have other applications beyond the space object tracking problem, such as in multi-target tracking with a group of autonomous robots [159], battlefield life signs detection by UAVs (Unmanned Aerial Vehicles) [160], package tracking in warehouse by sensor networks [161], etc.

5.2.2 Network Model

We define the *active node set* S_k^j in a time-varying graph $G(j)$ as the set of nodes that have updated local estimates to be shared with others in the j-th consensus iteration after the k-th tracking

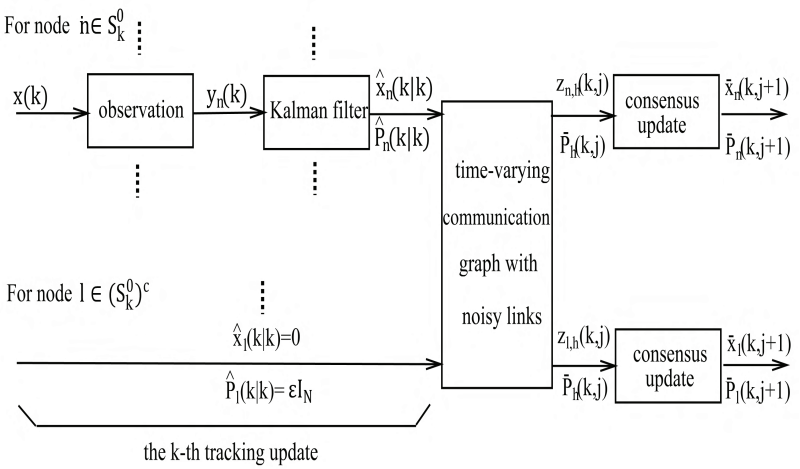


Figure 5.1: Block diagram of distributed tracking with consensus on a time-varying graph with incomplete data and noisy communication links.

update [119]. Define *effective network graph* of a network $G(j) = \big(V(j), E(j)\big)$ with active node set S_k^j as $\tilde{G}(j) = \big(V(j), \tilde{E}(j)\big)$, where $\tilde{E}(j) = E(j) \cap \big(\cup_{n \in S_k^j} \Upsilon_n^{\text{out}}(j)\big)$, where $\Upsilon_n^{\text{out}}(j) = \{(n,l) | (n,l) \in E(j)\}$ denotes the set of directed edges with initial vertex as n at iteration j. Note that, the effective network graph $\tilde{G}(j)$ is a directed graph, which is obtained by removing the outgoing edges of the nodes that do not have data in $G(j)$. For a static graph $G(j) = G(V,E)$, $\tilde{E}(j)$ can be written as $\tilde{E}(j) = \tilde{E}(j-1) \cup_{l \in S_k^{j-1}} \big(\cup_{n \in \Omega_l} \Upsilon_n^{\text{out}}\big)$, where $\tilde{E}(0) = E \cap \big(\cup_{n \in S_k^0} \Upsilon_n^{\text{out}}\big)$. Note that, the nodes that do not observe the target will not have updated local estimates to share at the beginning of consensus update process (at $j = 0$). However, as information exchange among nodes progresses, some of these nodes may be able to form their own updated local estimates at the consensus iteration j for $j > 0$. Therefore, the active node set S_k^j is time-varying and $S_k^j = S_k^{j-1} \cup_{l \in S_k^{j-1}} \Omega_l(j-1)$, where S_k^0 is the set of nodes that have observations of the target in the k-th tracking update step as in Fig. 5.1. Figure shows the relation between the connectivity graph $G(j)$ and the effective network graph $\tilde{G}(j)$ for a graph of 6 nodes with active node set $S_k^j = (1,2,4,6)$, where solid circles denote active nodes.

Let $I_{S_k^j}$ denote an $N \times N$ matrix generated from the active node set S_k^j as follows:

$$[I_{S_k^j}]_{nn'} = \begin{cases} 1 & \text{if } n = n' \quad \text{and} \quad n' \in S_k^j \\ 0 & \text{else} \end{cases} .$$

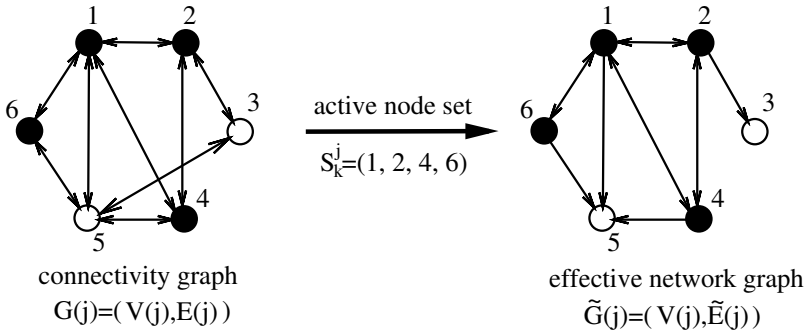


Figure 5.2: Connectivity graph and effective network graph.

Note that, $I_{S_k^j}$ is a diagonal matrix with n'-th diagonal element equal to zero for $n' \in (S_k^j)^c$, where $(\cdot)^c$ denotes the set complement. By combining the connectivity graph $G(j)$ with the active node set S_k^j, we obtain the effective network graph $\tilde{G}(j)$. Thus, the adjacency matrix of the effective network graph is given by $\mathtt{A}(j) = A(j)I_{S_k^j}$. The corresponding degree matrix $\mathtt{D}(j)$ can then be obtained from $\mathtt{A}(j)$, and the Laplacian matrix is $\mathtt{L}(j) = \mathtt{D}(j) - \mathtt{A}(j)$ by definition.

As an example, consider the same network model in Fig. 5.2. The matrix $I_{S_k^j} = \text{diag}(1, 1, 0, 1, 0, 1)$. The Laplacian matrices of the connectivity graph and effective network graph are as follows:

$$L(j) = \begin{bmatrix} 4 & -1 & 0 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 & -1 & 0 \\ -1 & -1 & 0 & 3 & -1 & 0 \\ -1 & 0 & -1 & -1 & 4 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathtt{L}(j) = \begin{bmatrix} 3 & -1 & 0 & -1 & 0 & -1 \\ -1 & 2 & 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & -1 & 3 & -1 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

5.2.3 Distributed Tracking with Consensus Algorithm

In this subsection, we propose a distributed tracking and consensus algorithm for the above distributed tracking problem over a time-varying graph with incomplete data and noisy communication links. This algorithm is based on the architecture that was first proposed in [119] in the special context of consensus tracking in a satellite sensor network for situational awareness.

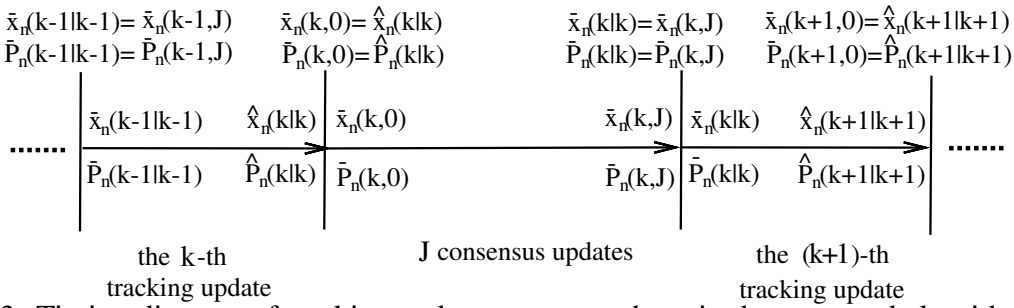


Figure 5.3: Timing diagram of tracking and consensus updates in the proposed algorithm for distributed tracking with consensus.

Figure 5.3 shows the timing diagram of tracking and consensus updates process in the proposed distributed tracking with consensus algorithm. As in Fig. 5.3, at tracking time step k, node n is assumed to have completed its consensus iterations corresponding to time $k-1$. If the output of this consensus update following the $(k-1)$-th tracking update step is $\overline{\mathbf{x}}_n(k-1,J)$ with the associated covariance matrix $\overline{P}_n(k-1,J)$, then node n sets $\overline{\mathbf{x}}_n(k-1|k-1)=\overline{\mathbf{x}}_n(k-1,J)$ and $\overline{P}_n(k-1|k-1)=\overline{P}_n(k-1,J)$. Next, at the k-th tracking update step, each node n where $n\in S_k^j$, passes its observation $\mathbf{y}_n(k)$ through its local Kalman filter as follows [91]:

$$\begin{aligned}
\hat{\mathbf{x}}_n(k|k-1) &= F\overline{\mathbf{x}}_n(k-1|k-1),\\
\hat{P}_n(k|k-1) &= F\overline{P}_n(k-1|k-1)F^T+Q,\\
K_n(k) &= \hat{P}_n(k|k-1)H_n^T\Big(H_n\hat{P}_n(k|k-1)H_n^T+R_n\Big)^{-1},\\
\hat{\mathbf{x}}_n(k|k) &= \hat{\mathbf{x}}_n(k|k-1)+K_n(k)\Big(\mathbf{y}_n(k)-H_n\hat{\mathbf{x}}_n(k|k-1)\Big),\\
\hat{P}_n(k|k) &= \Big(I-K_n(k)H_n\Big)\hat{P}_n(k|k-1),
\end{aligned}\tag{5.4}$$

where $\overline{\mathbf{x}}_n(k-1|k-1)=\overline{\mathbf{x}}_n(k-1,J)$ with $\overline{\mathbf{x}}_n(0,J)=\overline{\mathbf{x}}(0)$ and $\overline{P}_n(k-1|k-1)=\overline{P}_n(k-1,J)$ with $\overline{P}_n(0,J)=P_0$. Let $\overline{\mathbf{X}}(k-1,j)=\big[\overline{\mathbf{x}}_1(k-1,j)^T,\,\overline{\mathbf{x}}_2(k-1,j)^T,\cdots,\overline{\mathbf{x}}_N(k-1,j)^T\big]^T$. Denote $\overline{P}(k-1,j)$ as the covariance matrix corresponding to $\overline{\mathbf{X}}(k-1,j)$. The $\overline{P}_n(k-1,J)$ in (5.4) can be obtained by extracting the n-th $M\times M$ main diagonal block of $\overline{P}(k-1,J)$.

Node n uses its filtered estimate $\hat{\mathbf{x}}_n(k|k)$ obtained by the above tracking update step as the initial estimate for consensus update exchanges by setting $\overline{\mathbf{x}}_n(k,0)=\hat{\mathbf{x}}_n(k|k)$ with initial covariance

matrix $\overline{P}(k,0) = \hat{P}_1(k|k) \oplus \hat{P}_2(k|k) \oplus \cdots \oplus \hat{P}_N(k|k)$[1], where $\oplus$ denotes the direct sum. On the other hand, for nodes $n \in (S_k^j)^c$, we may arbitrarily set $\hat{\mathbf{x}}_n(k|k) = \mathbf{0}$ and $\hat{P}_n(k|k) = \epsilon I_M$ for some $\epsilon > 0$.

During the $(j+1)$-th consensus update, each node n forms a linear estimate of the following form as its consensus estimate:

$$\overline{\mathbf{x}}_n(k,j+1) = \overline{\mathbf{x}}_n(k,j) + \gamma_n(j) \sum_{l=1}^{N} \mathbb{A}_{n,l}(j) \Big(\overline{\mathbf{z}}_{n,l}(k,j) - \overline{\mathbf{z}}_{n,n}(k,j) \Big), \tag{5.5}$$

where $\gamma_n(j)$ is the n-th node's weight coefficient at iteration j and $0 \leq j < J$. We set $\gamma_n(j) = \gamma(j)$ for $n \in S_k^j$ and $\gamma_n(j) = \frac{1}{\sum_{l=1}^{N} \mathbb{A}_{n,l}(j)}$ for $n \in (S_k^j)^c$ and $\sum_{l=1}^{N} \mathbb{A}_{n,l}(j) \neq 0$. For node n that does not have local tracking estimate, we assume that it will generate its estimate by averaging the tracking estimates from its neighbors[2].

By defining $\overline{\mathbf{X}}(k,j) = \big[\overline{\mathbf{x}}_1(k,j)^T,\ \overline{\mathbf{x}}_2(k,j)^T, \cdots, \overline{\mathbf{x}}_N(k,j)^T\big]^T$, the consensus update dynamics can be written in vector form as follows:

$$\overline{\mathbf{X}}(k,j+1) = \overline{\mathbf{X}}(k,j) - \Big[\big(\Gamma(j)\mathtt{L}(j)\big) \otimes I_M \Big] \overline{\mathbf{X}}(k,j) - \big(\Gamma(j) \otimes I_M\big) \overline{\Phi}(j), \tag{5.6}$$

where $\Gamma(j) = \text{diag}\big(\gamma_1(j), \cdots, \gamma_N(j)\big)$, $\overline{\Phi}(j) = \big[\boldsymbol{\phi}_1(j)^T \cdots \boldsymbol{\phi}_N(j)^T\big]^T$ and $\boldsymbol{\phi}_n(j) = -\sum_{l \in \Omega_n(j)} \boldsymbol{\phi}_{n,l}(j)$. Note that, from (5.3), $\mathbb{E}[\overline{\Phi}(j)] = \mathbf{0}$ and $\sup_j \mathbb{E}[\|\overline{\Phi}(j)\|^2] = \eta \leq N(N-1)u < \infty$.

Let us define $\overline{\mathbf{e}}(k,j)$ to be the error vector at the j-th consensus iteration after the k-th tracking update: $\overline{\mathbf{e}}(k,j) \triangleq \overline{\mathbf{X}}(k,j) - (\mathbf{1} \otimes I_M)\mathbf{x}(k)$. From (5.6), it follows that

$$\begin{aligned} \overline{\mathbf{e}}(k,j+1) = &\big(\mathbf{A}(j) \otimes I_M\big) \overline{\mathbf{e}}(k,j) - \big(\Gamma(j) \otimes I_M\big) \overline{\Phi}(j) \\ &+ \Big(\big(\mathbf{A}(j) \otimes I_M\big) - I \Big) (\mathbf{1} \otimes I_M)\mathbf{x}(k), \end{aligned} \tag{5.7}$$

where $\mathbf{A}(j) = I_N - \Gamma(j)\mathtt{L}(j)$. Note that, this coefficient matrix $\mathbf{A}(j)$ is slightly different from the one proposed in [119]. In [119], $\mathbf{A}(j) = \tilde{\mathbf{I}}(j) - \gamma(j)\tilde{L}(j)$, where $\tilde{\mathbf{I}}(j)$ and $\tilde{L}(j)$ are the modified

[1]The assumption here is that at the beginning of the consensus update process, the filtered estimates at different nodes are statistically uncorrelated.

[2]Note that, for $n \in (S_k^j)^c$ and $\sum_{l=1}^{N} \mathbb{A}_{n,l}(j) = 0$, node n does not receive information from any node that has local tracking estimate. Then, $\overline{\mathbf{x}}_n(k,j+1) = \overline{\mathbf{x}}_n(k,j)$.

identity and Laplacian matrices. The required modification however does not lend itself to a convenient relation between the original matrices and the modified ones that can be used in mathematical derivations.

Note that, if the filtered estimate $\hat{\mathbf{x}}_n(k|k)$ at the end of the measurement update stage is an unbiased estimate, then $\overline{\mathbf{x}}_n(k,0)$ is also unbiased for all $n \in S_k^j$. From (5.5), since $\overline{\mathbf{x}}_n(k,j+1) = \frac{1}{\sum_{l=1}^N \mathbb{A}_{n,l}(j)} \sum_{l=1}^N \mathbb{A}_{n,l}(j) \left(\overline{\mathbf{x}}_l(k,j) + \boldsymbol{\phi}_{n,l}(j)\right)$ for $n \in (S_k^j)^c$, then $\overline{\mathbf{x}}_n(k,j+1)$ is also unbiased for $n \in (S_k^j)^c$ if $\overline{\mathbf{x}}_l(k,j)$ is unbiased for $l \in S_k^j$. From (5.7), it can be shown that the unbiasedness in consensus estimate $\overline{\mathbf{X}}(k,j)$ can be maintained if matrix $\mathbf{A}(j)$ satisfies the condition $\Big(\big(\mathbf{A}(j) \otimes I_M\big) - I\Big)(\mathbf{1} \otimes I_M) = \mathbf{0}$, which is equivalent to requiring $\Big(\big(\mathbf{A}(j) - I_N\big)\mathbf{1}\Big) \otimes I_M = \mathbf{0}$. It follows that the unbiasedness in consensus estimate $\overline{\mathbf{X}}(k,j)$ requires 0 to be an eigenvalue of the Laplacian matrix $\mathtt{L}(j)$ with the associated eigenvector $\mathbf{1}$[3]. Define the covariance matrix corresponding to $\overline{\mathbf{X}}(k,j)$ as $\overline{P}(k,j) = \mathbb{E}[\overline{\mathbf{e}}(k,j)\overline{\mathbf{e}}(k,j)^T]$. From (5.7) and unbiasedness condition, it can be easily seen that

$$\overline{P}(k,j+1) = \big(\mathbf{A}(j) \otimes I_M\big)\overline{P}(k,j)\big(\mathbf{A}(j) \otimes I_M\big)^T + \mathbb{E}\Big\{\big(\Gamma(j) \otimes I_M\big)\overline{\Phi}(j)\overline{\Phi}(j)^T\big(\Gamma(j) \otimes I_M\big)^T\Big\}. \quad (5.8)$$

As shown in Fig. 5.3, after J consensus iterations each node n will feed $\overline{\mathbf{x}}_n(k,J)$ back to their local Kalman filters by setting $\overline{\mathbf{x}}_n(k|k) = \overline{\mathbf{x}}_n(k,J)$ and $\overline{P}_n(k|k) = \overline{P}_n(k,J)$ before starting next tracking update at time $k+1$. Recall that here $\overline{P}_n(k,J)$ is the n-th $M \times M$ main diagonal block of $\overline{P}(k,J)$. Algorithm 1 shows a summary of the steps in the proposed distributed tracking with consensus algorithm.

[3]Note that, similar results on the unbiasedness of consensus estimate were obtained in [119].

5.3 Performance Analysis

5.3.1 Conditions for Achieving Consensus

In this section, we analyze the convergence of the proposed distributed tracking with consensus algorithm and the convergence rate. Note that, the proofs of lemmas and theorems in this section are different from those in [107] due to vector state and incomplete data, which results in two stages of consensus process: obtaining complete data from incomplete data and reaching consensus on complete data. In the scenarios we consider, we assume that the information exchange rate during the consensus update process is much higher compared to the data sampling rate for the tracking updates. Hence, we can assume that $J \gg 1$[4], guaranteeing enough time for information to be exchanged over the network so that consensus can be reached if the weight $\{\gamma(j)\}$ is chosen properly. As mentioned above, for a fixed k and $J \gg 1$, the consensus update process after the k-th tracking update can be considered as a consensus in estimation problem. Thus, to simplify notation, in the following we omit the tracking time step index k in $\overline{\mathbf{X}}(k, j)$.

We start by defining the consensus subspace $\mathcal{C}$ as

$$\mathcal{C} = \left\{ \mathbf{X} \in \mathbb{R}^{NM} | \mathbf{X} = \mathbf{1}_N \otimes \mathbf{a}, \mathbf{a} \in \mathbb{R}^M \right\}.$$

If the consensus algorithm (5.6) converges to the consensus subspace $\mathcal{C}$, each node estimate $\overline{\mathbf{x}}_n(j)$ will converge to the same value $\overline{\mathbf{x}}_n(j) = \mathbf{a}$ for $1 \leq n \leq N$, $\mathbf{a} \in \mathbb{R}^M$ and consensus is reached over the network. It is well known from the stochastic approximation literature [164] that, in order to ensure asymptotic convergence to consensus subspace, the weight coefficient $\gamma(j)$ must satisfy the *persistence condition* as follows

$$\gamma(j) > 0, \quad \sum_{j=0}^{\infty} \gamma(j) = \infty, \quad \sum_{j=0}^{\infty} \gamma(j)^2 < \infty. \tag{5.9}$$

[4]For practical consideration, due to energy constraints of sensor networks, the time period J for consensus process is not too long such that the nodes can still efficiently obtain new information from the source [162]. Simulation results in Section IV show how the algorithm performs in this case.

We recall the following result on distance properties in $\mathbb{R}^{NM}$:

Lemma 5.3.1. *Suppose that $\boldsymbol{X} \in \mathbb{R}^{NM}$ and consider the orthogonal decomposition $\boldsymbol{X} = \boldsymbol{X}_{\mathcal{C}} + \boldsymbol{X}_{\mathcal{C}^\perp}$. Then the Euclidean distance $\rho(\boldsymbol{X}, \mathcal{C}) = \|\boldsymbol{X}_{\mathcal{C}^\perp}\|$.*

In the following, we prove that the consensus algorithm given in (5.6) converges almost surely (a.s.). This is achieved in two steps: First, Lemma 5.3.3 proves that the state vector sequence $\{\overline{\mathbf{X}}(j)\}_{j\geq 0}$ converges a.s. to the consensus subspace $\mathcal{C}$. Theorem 5.3.4 then completes the proof by showing that the sequence of component-wise averages $\{\overline{\mathbf{X}}_{\text{avg}}(j)\}_{j\geq 0}$ converges a.s. to a finite random variable $\boldsymbol{\Theta}$, where $\overline{\mathbf{X}}_{\text{avg}}(j) = \frac{1}{N}(\mathbf{1}^T \otimes I_M)\overline{\mathbf{X}}(j)$. The proof of Theorem 5.3.4 will require a basic result on convergence of Markov processes from [164], which is restated as Lemma 5.3.2 in our context. Before stating the lemma, however, we need to introduce the notation of [164].

Let$\{\overline{\mathbf{X}}(j)\}_{j\geq 0}$ be a Markov process in $\mathbb{R}^{NM}$. Define the generating operator $\mathcal{L}$ corresponding to $\{\overline{\mathbf{X}}(j)\}_{j\geq 0}$ as

$$\mathcal{L}V(j, \overline{\mathbf{X}}) = \mathbb{E}\Big\{V\big(j+1, \overline{\mathbf{X}}(j+1)\big)|\overline{\mathbf{X}}(j) = \overline{\mathbf{X}}\Big\} - V(j, \overline{\mathbf{X}}),$$

for functions $V(j, \overline{\mathbf{X}})$, $j \geq 0$, $\overline{\mathbf{X}} \in \mathbb{R}^{NM}$, provided the conditional expectation exists. If $D_{\mathcal{L}}$ is the domain of $\mathcal{L}$, then we say that $V(j, \overline{\mathbf{X}}) \in D_{\mathcal{L}}$ in a domain $\mathcal{C}$, if $\mathcal{L}V(j, \overline{\mathbf{X}})$ is finite for all $(j, \overline{\mathbf{X}}) \in \mathcal{C}$.

For $G \subset \mathbb{R}^{NM}$, the ϵ-neighborhood of G and its complement are defined as,

$$U_\epsilon(G) = \left\{\mathbf{X}| \inf_{\mathbf{Y}\in G} \rho(\mathbf{X}, \mathbf{Y}) < \epsilon\right\}, \qquad V_\epsilon(G) = \mathbb{R}^{NM} \setminus U_\epsilon(G). \tag{5.10}$$

With these notations, we may now state the desired lemma on the convergence of Markov processes:

Lemma 5.3.2. *(Convergence of Markov Processes): Let $\{\overline{\boldsymbol{X}}(j)\}_{j\geq 0}$ be a Markov process with generating operator $\mathcal{L}$. Let there exist a non-negative function $V(j, \overline{\boldsymbol{X}}) \in D_{\mathcal{L}}$ in the domain*

$G \subset \mathbb{R}^{NM}$ for $j \geq 0$ and $\overline{\boldsymbol{X}} \in \mathbb{R}^{NM}$. Assume that

$$\inf_{j\geq 0, \overline{\boldsymbol{X}} \in V_\epsilon(G)} V(j,\overline{\boldsymbol{X}}) > 0, \quad \forall \epsilon > 0, \qquad \textit{and} \qquad V(j,\overline{\boldsymbol{X}}) = 0, \quad \overline{\boldsymbol{X}} \in G,$$

$$\lim_{\overline{\boldsymbol{X}} \to G} \sup_{j\geq 0} V(j,\overline{\boldsymbol{X}}) = 0, \qquad \textit{and} \qquad \mathcal{L}V(j,\overline{\boldsymbol{X}}) \leq g(j)\big(1 + V(j,\overline{\boldsymbol{X}})\big) - \gamma(j)\varphi(j,\overline{\boldsymbol{X}}),$$

where $\varphi(j,\overline{\boldsymbol{X}}), \overline{\boldsymbol{X}} \in \mathbb{R}^{NM}$ is a non-negative function such that

$$\inf_{j, \overline{\boldsymbol{X}} \in V_\epsilon(G)} \varphi(j,\overline{\boldsymbol{X}}) > 0, \quad \forall \epsilon > 0; \gamma(j) > 0, \quad \sum_{j\geq 0} \gamma(j) = \infty; \quad \textit{and} \quad g(j) > 0, \quad \sum_{j\geq 0} g(j) < \infty.$$

Then the Markov process $\big\{\overline{\boldsymbol{X}}(j)\big\}_{j\geq 0}$ with an arbitrary initial distribution converges almost surely (a.s.) to G as $j \to \infty$:

$$\mathbb{P}\left(\lim_{j\to\infty} \rho\left(\overline{\boldsymbol{X}}(j), G\right) = 0\right) = 1.$$

Proof. Proof is a vector generalization of that in [107], and is omitted. □

Lemma 5.3.2 guarantees a.s. convergence of a general Markov process with an arbitrary initial distribution under the assumption of the existence of a Lyapunov function $V(j,\overline{\mathbf{X}})$. In fact, the state vector sequence $\big\{\overline{\mathbf{X}}(j)\big\}_{j\geq 0}$ given in (5.6) is a Markov process, since $\mathbb{P}\big[\overline{\mathbf{X}}(j)|\overline{\mathbf{X}}(j-1), \cdots, \overline{\mathbf{X}}(0)\big] = \mathbb{P}\big[\overline{\mathbf{X}}(j)|\overline{\mathbf{X}}(j-1)\big]$. In the next lemma, we prove that the state estimate sequence $\big\{\overline{\mathbf{X}}(j)\big\}_{j\geq 0}$ given in (5.6) converges a.s. to the consensus subspace $\mathcal{C}$ by showing that the consensus algorithm over an undirected effective network graph satisfies the Lyapunov function assumptions of Lemma 5.3.2 .

Lemma 5.3.3. *(a.s. convergence of the proposed algorithm to the consensus subspace): Consider the consensus algorithm in (5.6) with initial state $\overline{\boldsymbol{X}}(0) \in \mathbb{R}^{NM}$. The weight coefficients satisfy the persistence condition in (5.9). Assume that the undirected connectivity graph Laplacian $L(j)$ is independent of communication noise $\phi_{n,l}(j)$ for $1 \leq n, l \leq N$. If $L(j) = \overline{L} + \tilde{L}(j)$ with mean $\overline{L} = \mathbb{E}\big[L(j)\big]$ is such that $\lambda_2(\overline{L}) > 0$ and $p(l,n) > 0$ for $\{l,n\} \in E(j)$, then*

$$\mathbb{P}\left[\lim_{j\to\infty} \rho\left(\overline{\boldsymbol{X}}(j), \mathcal{C}\right) = 0\right] = 1.$$

Proof. See Appendix B. □

Lemma 5.3.3 shows that the state estimate sequence $\{\overline{\mathbf{X}}(j)\}_{j\geq 0}$ given in (5.6) converges a.s. to the consensus subspace $\mathcal{C}$. The key to the proof is to show that the directed effective network graph will become an undirected graph after all nodes have local estimates and the consensus algorithm over this undirected effective network graph satisfies the condition required in Lemma 5.3.2. In the following theorem, we state our main result and complete the convergence proof for the proposed distributed tracking with consensus algorithm by showing that the sequence of component-wise averages $\{\overline{\mathbf{X}}_{\text{avg}}(j)\}_{j\geq 0}$ converges a.s. to a finite random variable $\boldsymbol{\Theta}$, where $\overline{\mathbf{X}}_{\text{avg}}(j) = \frac{1}{N}(\mathbf{1}^T \otimes I_M)\overline{\mathbf{X}}(j)$.

Theorem 5.3.4. *(a.s. convergence to a finite random vector): Consider the consensus algorithm in (5.6) with initial state $\overline{\boldsymbol{X}}(0) \in \mathbb{R}^{NM}$. The weight coefficients satisfy the persistence condition in (5.9). Assume that the time-varying connectivity graph Laplacian $L(j)$ is independent of communication noise $\phi_{n,l}(j)$ for $1 \leq n, l \leq N$. If $L(j) = \overline{L} + \tilde{L}(j)$ with mean $\overline{L} = \mathbb{E}\big[L(j)\big]$ is such that $\lambda_2(\overline{L}) > 0$, and if $p(l,n) > 0$ for $\{l,n\} \in E(j)$, then there exists an almost sure finite real random vector $\boldsymbol{\Theta}$ such that*

$$\mathbb{P}\left[\lim_{j\to\infty} \overline{\boldsymbol{X}}(j) = \boldsymbol{1}_N \otimes \boldsymbol{\Theta}\right] = 1.$$

Proof. Since the mean connectivity graph $\overline{L}$ is connected with non-zero link probability, for j large enough, each node will receive information from one another and generate its updated local estimate. For a fixed k, let $J_k = \inf\{j | (S_k^j)^c = \emptyset, j \geq 0\}$. Then, $\Gamma(j) = \gamma(j) I_N$ for $j \geq J_k$ and (5.6) becomes

$$\overline{\mathbf{X}}(j+1) = \overline{\mathbf{X}}(j) - \gamma(j)\Big[\big(\mathtt{L}(j) \otimes I_M\big)\overline{\mathbf{X}}(j) + \overline{\Phi}(j)\Big] \quad \text{for} \qquad j \geq J_k. \tag{5.11}$$

Define the average of $\overline{\mathbf{X}}(j)$ as $\overline{\mathbf{X}}_{\text{avg}}(j) = \frac{1}{N}(\mathbf{1}^T \otimes I_M)\overline{\mathbf{X}}(j)$. Multiply both sides of (5.11) by

$\frac{1}{N}(\mathbf{1}^T \otimes I_M)$ and use the fact that $\mathbf{1}^T \mathtt{L}(j) = \mathbf{0}_N$, so that for $(S_k^j)^c = \emptyset$. Then we have

$$\overline{\mathbf{X}}_{\text{avg}}(j+1) = \overline{\mathbf{X}}_{\text{avg}}(j) - \varepsilon(j) = \overline{\mathbf{X}}_{\text{avg}}(J_k) - \sum_{J_k \le l \le j} \varepsilon(l),$$

where $\varepsilon(j) = \frac{\gamma(j)}{N}(\mathbf{1}^T \otimes I_M)\overline{\Phi}(j)$. Assuming that receiver noise is zero-mean and time independent, we obtain

$$\mathbb{E}\big[\|\varepsilon(j)\|^2\big] = \frac{\gamma^2(j)}{N^2}\mathbb{E}\big[\overline{\Phi}(j)^T(\mathbf{1}^T \otimes I_M)^T(\mathbf{1}^T \otimes I_M)\overline{\Phi}(j)\big] = \frac{\gamma^2(j)}{N^2}\mathbb{E}\left[\sum_{1 \le n \le N} \big(\boldsymbol{\phi}_n(j)\big)^T \boldsymbol{\phi}_n(j)\right],$$

where $\boldsymbol{\phi}_n(j) = -\sum_{l \in \Omega_n(j)} \boldsymbol{\phi}_{n,l}(j)$ denotes the total incoming noise from node $l \in \Omega_n(j)$ to node n and the last step follows from the independence of $\boldsymbol{\phi}_l(j)$ and $\boldsymbol{\phi}_n(j)$. By assuming that $\mathbb{E}\big[\boldsymbol{\phi}_{l,n}(j)\boldsymbol{\phi}_{l,n}(j)^T\big] = \sigma^2 I_M$ for $1 \le l, n \le N$, we obtain

$$\mathbb{E}\big[\|\varepsilon(j)\|^2\big] \le \frac{\gamma^2(j)}{N^2} MN(N-1)\sigma^2 = \frac{\gamma^2(j)M(N-1)}{N}\sigma^2.$$

From independence of $\overline{\mathbf{X}}(j)$ and $\overline{\Phi}(j)$ and the independence of noise over time, we then have that

$$\mathbb{E}\big[\|\overline{\mathbf{X}}_{\text{avg}}(j+1)\|^2\big] \le \mathbb{E}\big[\overline{\mathbf{X}}_{\text{avg}}(J_k)^T \overline{\mathbf{X}}_{\text{avg}}(J_k)\big] + \sum_{l \ge J_k}^{j} \frac{\gamma^2(l)M(N-1)}{N}\sigma^2 \le \infty.$$

Denote $\overline{\mathbf{X}}_{\text{avg}}(j) = \big[\overline{X}_{\text{avg},1}(j) \cdots \overline{X}_{\text{avg},M}(j)\big]^T$. It can be easily seen that

$$\mathbb{E}\big[\big(\overline{X}_{\text{avg},m}(j+1)\big)^2\big] \le \mathbb{E}\big[\big(\overline{X}_{\text{avg},m}(J_k)\big)^2\big] + \sum_{l \ge J_k}^{j} \frac{\gamma^2(l)(N-1)}{N}\sigma^2 \le \infty.$$

Hence, the sequence $\big\{\overline{X}_{\text{avg},m}(j)\big\}$ is an L_2 bounded martingale and thus converges a.s. in L_2 to a finite random scalar θ. Define $\overline{\mathbf{X}}_m(j) = \big[\mathbf{e}_m^T\overline{\mathbf{x}}_1, \cdots, \mathbf{e}_m^T\overline{\mathbf{x}}_N\big]^T$. From the conclusion of Lemma 5.3.3, we have that $\mathbb{P}\left[\lim_{j\to\infty} \|\overline{\mathbf{X}}(j) - \mathbf{1}_N \otimes \overline{\mathbf{X}}_{\text{avg}}(j)\| = 0\right] = 1$, which implies that $\mathbb{P}\left[\lim_{j\to\infty} \|\overline{\mathbf{X}}_m(j) - \overline{X}_{\text{avg},m}(j)\mathbf{1}_N\| = 0\right] = 1$. Then, we obtain that $\mathbb{P}\left[\lim_{j\to\infty} \overline{\mathbf{X}}_m(j) = \theta\mathbf{1}_N\right] = 1$ and the theorem follows. □

Theorem 5.3.4 shows that the proposed distributed tracking with consensus algorithm will reach consensus almost surely and the consensus estimate $\lim_{j\to\infty} \overline{\mathbf{x}}(j)$ is a finite random vector $\boldsymbol{\Theta}$.

Since the consensus algorithm in (5.6) falls in the framework of stochastic approximation, we may also analyze the convergence rate of the consensus algorithm based on the ODE method (Ordinary Difference Equation) [166]. The next theorem characterizes an upper bound to the convergence rate of the proposed distributed tracking with consensus algorithm.

Theorem 5.3.5. *(convergence rate): Consider the consensus algorithm in (5.6) with initial state* $\overline{\boldsymbol{X}}(0) \in \mathbb{R}^{NM}$. *The weight coefficients satisfy the persistence condition in (5.9) and* $\gamma(j) \leq \frac{2}{\lambda_2(\overline{L})+\lambda_n(\overline{L})}$. *Assume that the time-varying connectivity graph Laplacian* $L(j)$ *is independent of communication noise* $\phi_{n,l}(j)$ *for* $1 \leq n, l \leq N$. *For* $j \geq J_k$, *the effective network graph Laplacian is* $\mathtt{L}(j) = \overline{\mathtt{L}} + \tilde{\mathtt{L}}(j)$ *with mean* $\overline{\mathtt{L}} = \mathbb{E}\big[\mathtt{L}(j)\big]$. *If the connectivity graph Laplacian* $L(j)$ *with mean* $\overline{L} = \mathbb{E}\big[L(j)\big]$ *is such that* $\lambda_2(\overline{L}) > 0$, *and if* $p(l,n) > 0$ *for* $\{l,n\} \in E(j)$, *the convergence rate*[5], *of the proposed consensus algorithm is bounded by* $-\lambda_2(\overline{\mathtt{L}})\left(\frac{1}{J-J_k}\sum_{J_k \leq j \leq J}\gamma(j)\right)$.

Proof. For a fixed i, let $J_k = \inf\{j | (S_k^j)^c = \emptyset, j \geq 0\}$. From the asymptotic unbiasedness of $\boldsymbol{\Theta}$, we have $\lim_{j\to\infty}\mathbb{E}\big[\overline{\mathbf{X}}(j)\big] = \mathbf{1}_N \otimes \mathbf{r}$, where $\mathbf{r} = \overline{\mathbf{X}}_{\text{avg}}(J_k)$. For $j \geq J_k$, define $\Xi(j) = I_{NM} - \gamma(j)(\overline{\mathtt{L}} \otimes I_M)$, where $\overline{\mathtt{L}} = \mathbb{E}\big[\mathtt{L}(j)\big]$. Using the fact that $\mathtt{L}(j)$ and $\overline{\mathbf{X}}(j)$ are independent, and $\mathbb{E}\big[\overline{\Phi}(j)\big] = \mathbf{0}_{NM}$, from (5.6), we have that

$$\mathbb{E}\big[\overline{\mathbf{X}}(j+1)\big] = \Xi(j)\mathbb{E}\big[\overline{\mathbf{X}}(j)\big] = \prod_{l=J_k}^{j}\Xi(l)\mathbb{E}\big[\overline{\mathbf{X}}(J_k)\big], \quad \forall j \geq J_k. \tag{5.12}$$

From the persistence condition $\gamma(j) > 0, \sum_{j\geq 0}\gamma(j) = \infty$ and $\sum_{j\geq 0}\gamma^2(j) \leq \infty$ [107], it follows that $\gamma(j) \to 0$. From the mixed-product property of Kronecker product $(A\otimes B)(C\otimes D) = AB \otimes CD$ and $\big(I_{NM} - \gamma(j)\overline{\mathtt{L}}\big)\mathbf{1}_N = \mathbf{1}_N$ [158], we have

$$\mathbf{1}_N \otimes \mathbf{r} = \Xi(j)\big(\mathbf{1}_N \otimes \mathbf{r}\big). \tag{5.13}$$

[5]Note that the convergence rate calculated here is for the period of $J_k \leq j \leq J$, where $J \gg 1$ is the number of consensus iterations. From persistence condition (9), $\lim_{j\to\infty}\gamma(j) = 0$. Then $\gamma(j)$ is very close to zero and the convergence speed can be assumed negligible for $j \geq J$ and J large enough [118, 165].

From (5.12) and (5.13), it can be shown that

$$\begin{aligned}\|\mathbb{E}\big[\overline{\mathbf{X}}(j)\big] - \mathbf{1}_N \otimes \mathbf{r}\| &\leq \prod_{J_k \leq l \leq j-1} \overline{\rho}\big(1 - \gamma(l)\overline{\mathtt{L}}\big)\|\mathbb{E}\big[\overline{\mathbf{X}}(J_k)\big] - \mathbf{1}_N \otimes \mathbf{r}\| \\ &= \prod_{J_k \leq l \leq j-1} \big(1 - \gamma(l)\lambda_2(\overline{\mathtt{L}})\big)\|\mathbb{E}\big[\overline{\mathbf{X}}(J_k)\big] - \mathbf{1}_N \otimes \mathbf{r}\|,\end{aligned}$$

where last step follows from Lemma 8 of [106] and $\overline{\rho}(\cdot)$ denotes the spectral radius of a matrix. From the assumption on weight coefficient $\gamma(j)$, we have $0 \leq \gamma(l)\lambda_2(\overline{\mathtt{L}}) \leq 1$. Since $1 - \alpha \leq e^{-\alpha}$ for $0 \leq \alpha \leq 1$, we then have that

$$\|\mathbb{E}\big[\overline{\mathbf{X}}(j)\big] - \mathbf{1}_N \otimes \mathbf{r}\| \leq \Big(e^{-\lambda_2(\overline{\mathtt{L}})\left(\sum_{J_k \leq l \leq j-1} \gamma(l)\right)}\Big)\|\mathbb{E}\big[\overline{\mathbf{X}}(J_k)\big] - \mathbf{1}_N \otimes \mathbf{r}\|. \tag{5.14}$$

Therefore, as $j \to J$, the convergence rate is bounded by $-\lambda_2(\overline{\mathtt{L}})\left(\frac{1}{J - J_k}\sum_{J_k \leq l \leq J} \gamma(l)\right)$, which depends on the algebraic connectivity $\lambda_2(\overline{\mathtt{L}})$ and the weights $\gamma(j)$, for $J_k \leq j \leq J$. □

Theorem 5.3.5 shows that the convergence rate of the proposed algorithm depends on the topology through the algebraic connectivity $\lambda_2(\overline{\mathtt{L}})$ of the effective network graph $\tilde{G}(j)$ and through weights $\gamma(j)$, for $j \geq J_k$. Since for $j \geq J_k$, $I_{S_k^j} = I$ and $L(j) = \mathtt{L}(j)$, we have $\overline{\mathtt{L}} = \mathbb{E}\big[\mathtt{L}(j)\big] = \mathbb{E}\big[L(j)\big]$. In (5.14), $\lambda_2(\overline{\mathtt{L}})$ is the algebraic connectivity of the mean Laplacian corresponding to the time-varying network graphs. For a static network, this reduces to the algebraic connectivity of the static Laplacian $\mathtt{L}$.

Since the consensus algorithm in (5.6) is iterative, whose energy consumption is proportional to the time necessary to achieve consensus and inversely proportional to transmit power. From [162, 163], for energy-constrained sensor networks, there exists a trade-off between convergence time which depends on network connectivity and the transmit power of each node necessary to establish the links with the desired reliability. Therefore, we can minimize the energy consumption for consensus process by optimizing transmit power, network topology and weights $\gamma(j)$.

5.3.2 Steady-State Analysis for Noiseless Graphs

In this section, we analyze the steady-state performance of the proposed distributed tracking with consensus algorithm. When the filter reaches steady-state, the error covariance matrix is time-invariant and the corresponding filter gain is constant. Therefore, finding the steady-state of the proposed algorithm will help understanding its asymptotic behavior, analyzing error covariance and filter design. From (5.8), it can be seen that the propagation of communication noise implies the non-existence of an upper bound to the covariance matrix. Therefore, the covariance matrix in the Kalman filter may not also converge and the filter may not reach steady-state. However, time-varying graph assumption does not affect the existence of steady-state. Since for $J \to \infty$, consensus is reached over the network and the outputs of the consensus update $\overline{\mathbf{X}}_n(k, J)$ and $\overline{P}(k, J)$ depend only on the inputs $\overline{\mathbf{X}}_n(k, 0)$ and $\overline{P}(k, 0)$ for complete data case with noiseless time-varying graphs (for incomplete data case with noiseless time-varying graphs, this property still holds for some special types of graphs). Hence, the combined system of distributed tracking with consensus can be transformed into a Kalman filter with time-invariant parameters. Therefore, steady-state can still be reached [91]. In the following, assuming noiseless time-varying graphs, we start with steady-state analysis for the case with complete data and then we extend the results to the case with incomplete data.

Complete Data with Noiseless Time-varying Graphs

Here we assume complete data, a scalar target state $x \in \mathbb{R}^1$ (for simplicity) and noiseless time-varying graphs, where the connectivity graph Laplacian $L(j)$ with mean $\overline{L} = \mathbb{E}[L(j)]$ is such that $\lambda_2(\overline{L}) > 0$, and $p(l, n) > 0$ for $\{l, n\} \in E(j)$. Note that, since a closed form equation for $\hat{P}_n(k+1|k)$ can not be easily obtained when the target state $x \in \mathbb{R}^M$ for $M > 1$, the following derivation would not apply to vector state.

From the result of Theorem 5.3.4 for scalar target state, it can be shown that $\lim_{J\to\infty} \overline{\mathbf{X}}(k, J) = \overline{X}_{\text{avg}}(k, 0)\mathbf{1}_N$, where $\overline{X}_{\text{avg}}(k, j) = \frac{1}{N}\mathbf{1}^T\overline{\mathbf{X}}(k, J)$. From the definition of $\overline{\mathbf{X}}(k, j)$ and $\overline{x}_n(k, 0) =$

$\hat{x}_n(k|k)$, we have for $1 \leq n \leq N$

$$\lim_{J\to\infty} \overline{x}_n(k, J) = \frac{1}{N}\sum_{n=1}^{N} \hat{x}_n(k|k). \tag{5.15}$$

With the assumptions above, the covariance matrix (5.8) in the $(j+1)$-th consensus iteration after the k-th tracking update simplifies to $\overline{P}(k, j+1) = \mathbf{A}(j)\overline{P}(k, j)\mathbf{A}(j)^T$. For complete data case, $\mathtt{L}(j) = L(j)$. Since $\mathbf{1}^T L(j) = 0$, from (5.7) we have $\mathbf{1}^T\mathbf{A}(j) = \mathbf{1}$. Then, we can obtain that

$$\mathbf{1}^T\overline{P}(k, j+1)\mathbf{1} = \mathbf{1}^T\overline{P}(k, j)\mathbf{1}. \tag{5.16}$$

By applying the result of Theorem 5.3.4, we have $\lim_{J\to\infty} \overline{P}(k, J) = (\overline{X}_{\text{avg}}(k, 0) - x(k))^2\mathbf{1}\mathbf{1}^T$. Since all the elements in $\lim_{J\to\infty} \overline{P}(k, J)$ are equal, from (5.16), it follows that

$$\lim_{J\to\infty} \overline{P}(k, J) = \frac{\mathbf{1}^T\overline{P}(k, 0)\mathbf{1}}{N^2}\mathbf{1}\mathbf{1}^T = \frac{\sum_{n=1}^{N} \hat{P}_n(k|k)}{N^2}\mathbf{1}\mathbf{1}^T. \tag{5.17}$$

Since $\overline{P}_n(k, J)$ is the n-th $M \times M$ main diagonal block of $\overline{P}(k, J)$, we have the covariance matrix for node n ($1 \leq n \leq N$) as below:

$$\lim_{J\to\infty} \overline{P}_n(k, J) = \frac{\sum_{n=1}^{N} \hat{P}_n(k|k)}{N^2}. \tag{5.18}$$

From (5.15) and (5.18), we have $\overline{x}_n(k, J) = \overline{x}_l(k, J)$ and $\overline{P}_n(k, J) = \overline{P}_l(k, J)$ for $J \to \infty$ and $1 \leq n, l \leq N$. Then, each node n sets $\overline{x}_n(k|k) = \overline{x}_n(k, J)$ and $\overline{P}_n(k|k) = \overline{P}_n(k, J)$. From (5.4), we have $\hat{x}_n(k+1|k) = \hat{x}_l(k+1|k)$ and $\hat{P}_n(k+1|k) = \hat{P}_l(k+1|k)$ for $1 \leq n, l \leq N$ and it follows that for $1 \leq n \leq N$

$$\begin{aligned}
\hat{x}_n(k+1|k) &= F\frac{1}{N}\sum_{q=1}^{N} \hat{x}_q(k|k-1) - F\frac{1}{N}\sum_{q=1}^{N} \Big[K_q(k)\big(y_q(k) - H_q\hat{x}_q(k|k-1)\big)\Big], \\
\hat{P}_n(k+1|k) &= Q + \frac{1}{N^2}\sum_{q=1}^{N} F\big(I - K_q(k)H_q\big)\hat{P}_q(k|k-1)F^T.
\end{aligned} \tag{5.19}$$

Let $\hat{x}_n(k+1|k) = \hat{x}(k+1|k)$ and $\hat{P}_n(k+1|k) = \hat{P}(k+1|k)$. Then, the combined system of distributed tracking with consensus can be transformed into a single Kalman filter as follows:

$$\begin{aligned}
\hat{x}(k+1|k) &= F\hat{x}(k|k-1) + \frac{F}{N}\sum_{n=1}^{N}\Big[K_n(k)\big(y_n(k) - H_n\hat{x}_n(k|k-1)\big)\Big], \\
K_n(k) &= \hat{P}(k|k-1)H_n^T\left[H_n\hat{P}(k|k-1)H_n^T + R_n\right]^{-1}, \\
\hat{P}(k+1|k) &= Q + \frac{1}{N^2}\sum_{n=1}^{N}\Big[F\hat{P}(k|k-1)F^T \\
&\quad -FK_n(k)\left(H_n\hat{P}(k|k-1)H_n^T + R_n\right)K_n(k)^TF^T\Big]. \qquad (5.20)
\end{aligned}$$

Theorem 5.3.6. *Consider the system dynamics in (5.1) and (5.2) and the Kalman filter in (5.20). Assume that the connectivity graph Laplacian $L(j)$ with mean $\overline{L} = \mathbb{E}[L(j)]$ is such that $\lambda_2(\overline{L}) > 0$, and $p(l,n) > 0$ for $\{l,n\} \in E(j)$. If the pair (F, H_n) is observable for $1 \le n \le N$, then the prediction covariance matrix $\hat{P}(k|k-1)$ converges to a constant matrix*

$$\lim_{k\to\infty}\hat{P}(k|k-1) = P,$$

where P is the unique definite solution of the discrete algebraic Riccati equation (DARE)

$$P = Q + \frac{1}{N^2}\sum_{n=1}^{N}\left[FPF^T - FPH_n^T\left(H_nPH_n^T + R_n\right)^{-1}H_nPF^T\right]. \qquad (5.21)$$

Proof. See Proof of Theorem 5.3.7. By setting $m = N$ and $\beta_n(k) = 1$ for $1 \le n \le N$, the Kalman filter in (5.24) can be reduced to the one in (5.20). Theorem 5.3.6 can be considered as a special case of Theorem 5.3.7. Thus, it can be proved in a similar manner. □

As a consequence of Theorem 5.3.6, the local Kalman filter gain converges to

$$\lim_{k\to\infty}K_n(k) = PH_n^T\left[H_nPH_n^T + R_n\right]^{-1}.$$

From (5.21), it can be seen that $\lim_{N\to\infty}P = Q$. i.e. as the size of the sensor network N increases, the steady-state covariance P, which in this case is a scalar, will decrease. This implies that if the

network size is large enough, asymptotically the tracking is noiseless and follows the target exactly. It is obvious that this result still holds for distributed local Kalman filtering with centralized fusion. However, the distributed tracking with consensus results in the same performance even if the graph is time-varying and it also improves the robustness and scalability due to consensus exchanges. For the assumed scalar case, for example, if $H_n = H$ and $R_n = R$ for $1 \leq n \leq N$, then we have $K = \frac{HP}{H^2P+R}$ and $P = \frac{-B+\sqrt{B^2+4H^2QR}}{2H^2}$, where $B = \left(1 - \frac{F^2}{N}\right) R - H^2Q$. This implies that for the same sensing model, each node will have the same Kalman gain K and prediction covariance P in the steady-state.

Incomplete Data with Noiseless Time-varying Graphs

Next, we assume incomplete data, a scalar target state $x \in \mathbb{R}^1$ and noiseless time-varying graphs, where the connectivity graph Laplacian $L(j)$ with mean $\overline{L} = \mathbb{E}[L(j)]$ is such that $\lambda_2(\overline{L}) > 0$, and $p(l,n) > 0$ for $\{l,n\} \in E(j)$. Furthermore, we assume that only m nodes can observe the target and without loss of any generality the index of those m nodes are ordered as $1, 2, \cdots, m$, where m is constant and $1 \leq m \leq N$. This implies that the active node set $S_k^0 = \{1, 2, \cdots, m\}$, which does not require further assumptions on the connectivity graph for consensus, since the graph is connected on average and the information can still propagate over the network even if only a fixed number of nodes have observation.

With the assumption of incomplete data and noiseless time-varying graphs, $\mathbf{1}^T \mathtt{L}(j) = 0$ for $J_k \leq j < J$. Then, (5.17) becomes

$$\lim_{J\to\infty} \overline{P}(k,J) = \frac{\mathbf{1}^T \overline{P}(k,J_k)\mathbf{1}}{N^2}\mathbf{1}\mathbf{1}^T = \frac{\mathbf{1}^T \left[\mathbf{A}_0^{J_k-1}\right] \overline{P}(k,0) \left[\mathbf{A}_0^{J_k-1}\right]^T \mathbf{1}}{N^2}\mathbf{1}\mathbf{1}^T = \frac{\sum_{n=1}^m \hat{P}_n(k|k)\beta_n^2(k)}{N^2}\mathbf{1}\mathbf{1}^T \quad (5.22)$$

where $\left[\mathbf{A}_0^{J_k-1}\right] = \mathbf{A}(J_k-1)\cdots\mathbf{A}(0)$ and $\beta_n(k) = \sum_{l=1}^N \left[\mathbf{A}_0^{J_k-1}\right]_{ln}$ is the n-th column sum of $\left[\mathbf{A}_0^{J_k-1}\right]$ that depends on time k. The last step of (5.22) follows from that $\hat{P}_n(k|k) = \epsilon$ for $m <$

$n \leq N$ and some $\epsilon > 0$. Then, as in previous subsection, we have for $1 \leq n \leq N$

$$\lim_{J\to\infty} \overline{P}_n(k,J) = \frac{\sum_{n=1}^{m} \hat{P}_n(k|k)\beta_n^2(k)}{N^2} \qquad \text{and} \qquad \lim_{J\to\infty} \overline{x}_n(k,J) = \frac{1}{N}\sum_{n=1}^{m} \hat{x}_n(k|k)\beta_n(k). \tag{5.23}$$

From (5.23), for $J \to \infty$, we have $\overline{x}_n(k,J) = \overline{x}_l(k,J)$ and $\overline{P}_n(k,J) = \overline{P}_l(k,J)$ for $1 \leq n,l \leq N$. By setting $\overline{x}_n(k|k) = \overline{x}_n(k,J)$ and $\overline{P}_n(k|k) = \overline{P}_n(k,J)$, from (5.4), we can obtain recursive update equations for $\hat{P}_n(k+1|k)$ and $\hat{x}_n(k+1|k)$. Furthermore, we also have $\hat{x}_n(k+1|k) = \hat{x}_l(k+1|k)$ and $\hat{P}_n(k+1|k) = \hat{P}_l(k+1|k)$ for $1 \leq n,l \leq N$. Let $\hat{x}_n(k+1|k) = \hat{x}(k+1|k)$ and $\hat{P}_n(k+1|k) = \hat{P}(k+1|k)$. Then, the combined system of distributed tracking with consensus can then be transformed into a single Kalman filter for node n $(1 \leq n \leq m)$ as below:

$$\begin{aligned}
\hat{x}(k+1|k) &= \frac{F}{N}\sum_{n=1}^{m} \hat{x}(k|k-1)\beta_n(k) + \frac{F}{N}\sum_{n=1}^{m}\Big[K_n(k)\big(y_n(k) - H_n\hat{x}(k|k-1)\big)\beta_n(k)\Big], \\
K_n(k) &= \hat{P}(k|k-1)H_n^T\Big[H_n\hat{P}(k|k-1)H_n^T + R_n\Big]^{-1}, \\
\hat{P}(k+1|k) &= Q + \frac{1}{N^2}\sum_{n=1}^{m}\Big[F\hat{P}(k|k-1)F^T \\
&\qquad -FK_n(k)\Big(H_n\hat{P}(k|k-1)H_n^T + R_n\Big)K_n(k)^TF^T\Big]\beta_n^2(k),
\end{aligned} \tag{5.24}$$

where (5.20) is a special case of (5.24) with $m = N$ and $\beta_n(k) = 1$ for $1 \leq n \leq N$.

Theorem 5.3.7. *Consider the system dynamics in (5.1) and (5.2) and the Kalman filter in (5.24). Assume that m nodes can observe the target and the index of those m nodes are fixed and ordered as $1, \cdots, m$. The connectivity graph Laplacian $L(j)$ with mean $\overline{L} = \mathbb{E}[L(j)]$ is such that $\lambda_2(\overline{L}) > 0$, and $p(l,n) > 0$ for $\{l,n\} \in E(j)$. The connectivity graph has switching topologies and is periodic such that $\beta_n(k) = \beta_n$ is time-invariant. If the pair (F, H_n) is observable for $1 \leq n \leq m$, then the prediction covariance matrix $\hat{P}(k|k-1)$ converges to a constant matrix*

$$\lim_{k\to\infty} \hat{P}(k|k-1) = P,$$

where P is the unique definite solution of the discrete algebraic Riccati equation (DARE)

$$P = Q + \frac{1}{N^2}\sum_{n=1}^{m}\left[FPF^T - FPH_n^T\left(H_nPH_n^T + R_n\right)^{-1}H_nPF^T\right]\beta_n^2. \tag{5.25}$$

Proof. See Appendix C. □

Theorem 5.3.7 asserts that if the connectivity graph topology is switching and periodic, the proposed algorithm can still reach steady-state and the steady-state covariance matrix can be obtained by solving (5.25). The conditions of graph topology assumed in Theorem 5.3.7 are strong. However, it may still be applicable in certain situations such as satellite surveillance network in [119], since the existence of a communication link depends on distance between nodes and the trajectories of satellites are pre-determined and periodic, whenever ratios of the orbit periods are rational. As an example, consider the network model in Fig. 5.8. The connectivity graph in Fig. 5.8 is switching and periodic with period equal to 4 and it can be seen that the graph is connected on average. Let $N = 6$, $m = 4$, $S_k^0 = \{1, 2, 3, 4\}$ and $\gamma(j) = \frac{1}{j+1}$ for $0 \leq j < J$. After iteration $J_k = 1$, all nodes will have updated local estimates to be shared. In this case, $\left[\mathbf{A}_0^{J_k-1}\right]$ becomes

$$\left[\mathbf{A}_0^{J_k-1}\right] = \begin{bmatrix} -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

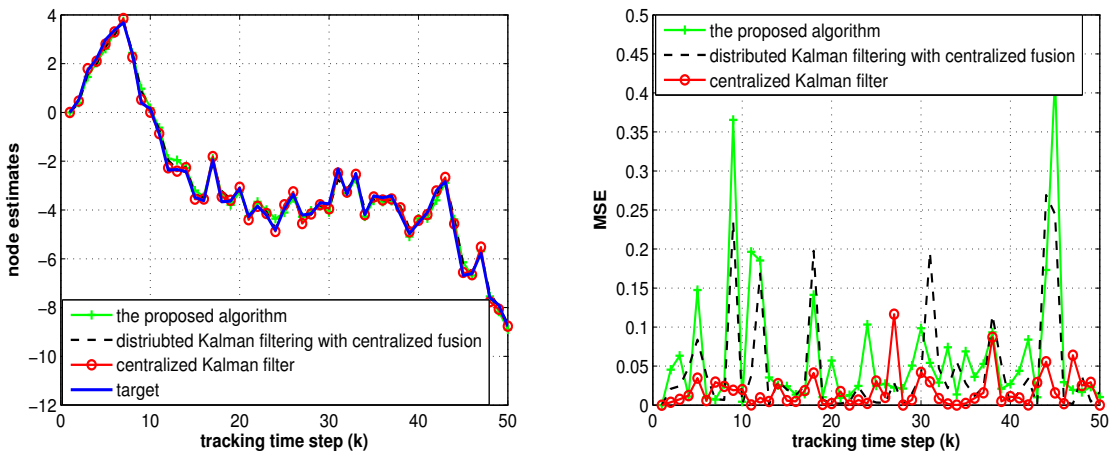
It can be seen that $\left[\mathbf{A}_0^{J_k-1}\right]$ is time-invariant, due to the time-variance of m and the periodic graph topology. Thus, $\beta_n(k) = \beta_n$ is also time-invariant and $\frac{1}{N}\sum_{n=1}^{m}\beta_n = 1$, which follows from the condition for unbiasedness in the consensus estimate $\overline{\mathbf{X}}(k, j)$. As we will see shortly, from the simulation results in Section IV the filter indeed reaches steady-state in this case, and then the error covariance matrix becomes time-invariant and the corresponding filter gain is constant.

5.4 Numerical Examples

In this section, we consider the performance of the proposed distributed tracking with consensus algorithm and compare it with centralized Kalman filter and distributed local Kalman filtering with centralized fusion. The performance of the centralized Kalman filter is well-understood [167] and provides a benchmark performance for distributed local Kalman filtering with centralized fusion. In distributed local Kalman filtering with centralized fusion, all nodes send their filtered estimates to a fusion center. The fusion center then generates a fused estimate $\hat{\mathbf{x}}_{\text{fusion}}(k) = \frac{1}{|S_k^0|}\sum_{n\in S_k^0}\hat{\mathbf{x}}_n(k|k)$.

In the first simulation we compare the performance of the proposed algorithm with the distributed local Kalman filtering with centralized fusion and the centralized Kalman filter over a random graph with noisy communication links and incomplete data. We consider a random connectivity graph $G(N,p)$ with $N = 20$ and the probability that each link exists $p = 0.5$. The other parameters of the simulation setup are: $F = 1$, $Q = 1$, $x(0) = 0$, $P_0 = 0$, $R_n = 0.25$, $H_n = 1$, $\sigma_{l,n}^2 = \sigma^2 = 0.1$, $S_k^0 = \{n|1 \leq n \leq 10, n \in \mathbb{Z}\}$ and $J = 30$.

Figure 5.4(a) shows the node estimates of the three algorithms in a time-varying graph with noisy communication links. As we can see, the node estimates of the three algorithms follow the target's trajectory. In Fig. 5.4(a), the curve with cross marker denotes the first node's estimate by using distributed tracking with consensus algorithm, the dashed curve denotes the distributed local Kalman filtering with centralized fusion, the curve with circle marker denotes the centralized Kalman filter and the solid curve denotes the target's trajectory. Figure 5.4(b) compares the resulting mean squared error (MSE) of the three algorithms, where the MSE of the distributed tracking with consensus is defined to be the average MSE over all nodes $\frac{1}{N}\sum_{n=1}^{N}\Big[\big(\overline{\mathbf{x}}_n(k,J)-\mathbf{x}(k)\big)^T \big(\overline{\mathbf{x}}_n(k,J)-\mathbf{x}(k)\big)\Big]$. In Fig. 5.4(b), it can be seen that the MSE of the proposed distributed tracking with consensus algorithm is close to that of the distributed local Kalman filtering with centralized fusion. As expected, both of them are higher than the MSE of the centralized Kalman filter, which acts as a benchmark. The results in Fig. 5.4 show that the performance of the pro-

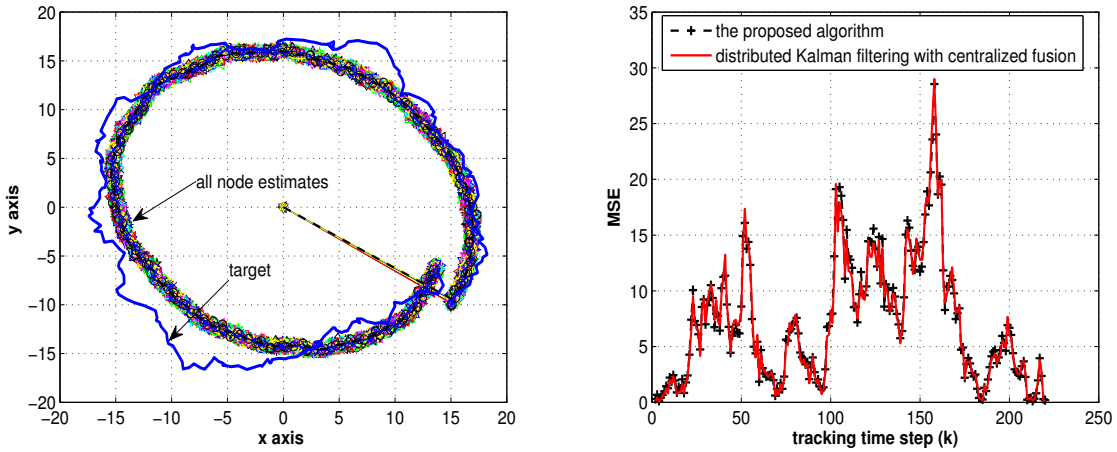


(a) (b)

Figure 5.4: Comparison of the proposed distributed tracking with consensus algorithm with distributed local Kalman filtering with centralized fusion and centralized Kalman filter – (a) node estimates, (b) mean squared error.

posed distributed tracking with consensus algorithm is close to that of the distributed local Kalman filtering with centralized fusion in a time-varying random graph with noisy communication and incomplete data. Additional communication bandwidth, which depends on graph topology G and number of iterations J, is required for the proposed algorithm due to information exchange among nodes. However, it resolves the bandwidth-constraints problem of fusion center for centralized fusion case and has a high level of fault tolerance and reliability. Also, because of its advantages of fully distributed implementation, robustness and scalability, it may be preferable in practical applications.

In the second simulation, we consider the two dimensional tracking problem treated in [114]. The connectivity graph is again assumed to be a random graph $G(N,p)$ with $N = 50$ and the probability that each link exists $p = 0.5$. The probability of each node having an observation at a given time instant is $p_s = 0.9$. The other parameters of the simulation setup are as follows: $F = I_2 + \epsilon F_0 + \frac{\epsilon^2}{2}F_0^2 + \frac{\epsilon^3}{6}F_0^3$, $F_0 = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$, $\epsilon = 0.015$, $Q = (\epsilon c_w^2)^2 I_2$, $c_w = 5$, $x(0) = [15, -10]^T$,



(a) (b)

Figure 5.5: Comparison of the proposed distributed tracking with consensus algorithm and distributed Kalman filter with centralized fusion in a two dimensional tracking problem – (a) trajectory, (b) mean squared error.

$H_n = [1, 0]$ for n is odd and $H_n = [0, 1]$ for n is even, $R_n = c_v^2\sqrt{n}$ for $n = 1, \cdots, N$ with $c_v = 30$, $\sigma_{l,n}^2 = \sigma^2 = 1$, $J = 10$. Note that, the target is moving on noisy circular trajectories. The target is not fully observable by an individual node, but is collectively observable by all nodes.

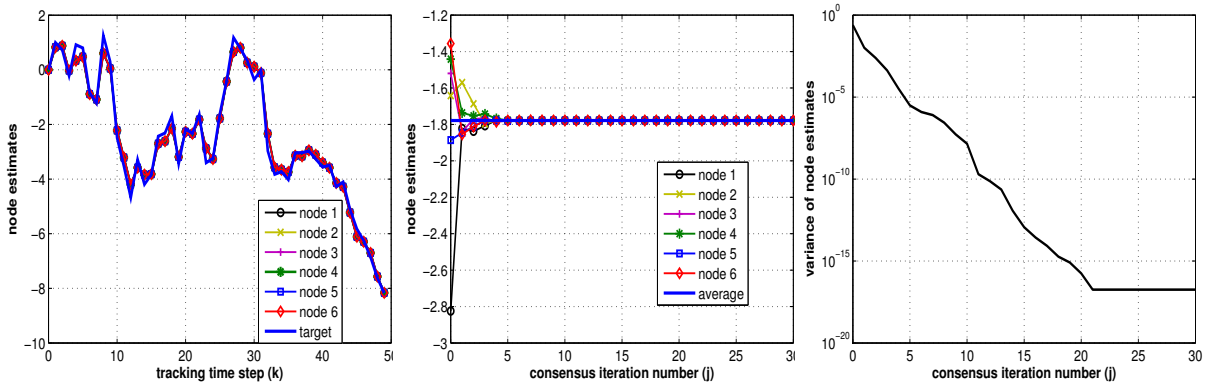
Figure 5.5(a) shows the node estimates (trajectory) of the two algorithms over a time-varying graph with incomplete data. In Fig. 5.5(a), the curves with markers denote all the node estimates by using distributed tracking with consensus algorithm, while the dashed curve denotes the distributed local Kalman filtering with centralized fusion and the solid curve denotes the target's trajectory. As we can see, both algorithms overcome the impact of partial observations at each node resulting in improved overall observation quality and the node estimates by using distributed tracking with consensus algorithm are noisy due to the communication noise. Note that the estimates are close to the trajectory of the target but with a small gap. That is because the observation noise covariance is rather large at each node. Figure 5.5(b) compares the resulting MSE of these algorithms. It can be seen that the mean squared error of the proposed algorithm is slightly higher than that of the distributed Kalman filtering with centralized fusion.

Next, we study the steady-state behavior in the case of time-varying graphs with complete data and noiseless communication. We consider a random connectivity graph $G(N,p)$ with $N=6$ and the probability that each link exists $p=0.5$. The other parameters of the simulation setup are as follows: $F=1$, $Q=1$, $x(0)=0$, $P_0=0.5$, $R_n=0.25$, $\sigma^2_{l,n}=\sigma^2=0$, $J=30$, $H_n=1$. Figure 5.6(a) shows the node consensus estimates $\overline{\mathbf{x}}_n(k,J)$ over a random graph with noiseless communication links and complete data. It can be seen that all node estimates $\overline{\mathbf{x}}_n(k,J)$ converge to the same value and follow the target state, as asserted by Theorem 5.3.4. Figure 5.6(b) and 5.6(c) show the node estimates $\overline{\mathbf{x}}_n(k,j)$ in the consensus update after the 21-st tracking update and the variance of all the node estimates, respectively. Here the variance of all the node estimates is defined as $var(k,j)=\mathbb{E}\left[\left(\overline{\mathbf{x}}_n(k,j)-\mu(k,j)\right)^T\left(\overline{\mathbf{x}}_n(k,j)-\mu(k,j)\right)\right]$, where $\mu(k,j)=\frac{1}{N}\sum_{n=1}^{N}\overline{\mathbf{x}}_n(k,j)$. From Fig. 5.6(b), it can be seen that the node estimates converge to the average which is also confirmed in Fig. 5.6(c), where the variance $var(k,j)$ decreases as consensus iteration number increases and becomes static (around 10^{-17}) after consensus is reached. Figure 5.7 shows the node estimate variance $\hat{P}_n(k|k-1)$ and Kalman gain $K_n(k)$ of the filter in (5.20). It can be seen that as the Kalman filter reaches steady-state, both the node estimate variance and the Kalman gain converge, as asserted by Theorem 5.3.6.

Next, we study the steady-state behavior on a graph with switching topologies and incomplete data and noiseless communication. The assumed parameters in the first simulation setup are as follows: $F=1$, $Q=1$, $x(0)=0$, $P_0=0.5$, $R_n=0.25$, $\sigma^2_{l,n}=\sigma^2=0$, $N=6$, $J=40$, $S^0_k=\{1,3,4,6\}$, $H_n=0.5$ for $n=1,3$ and $H_n=1$ for $n=4,6$. The connectivity graph Laplacian is

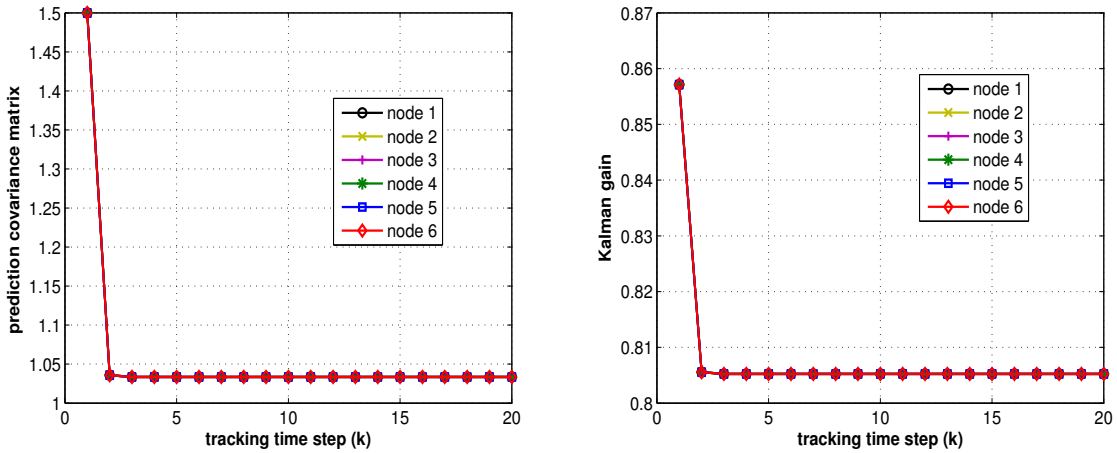
$$L(j)=\begin{cases} L_1 & j=4m \\ L_2 & j=4m+1 \\ L_3 & j=4m+2 \\ L_4 & j=4m+3 \end{cases} \quad \text{for} \quad m=0,1,2,\cdots,$$

which is shown in Fig. 5.8. As we can see, the graph is connected on average and $p(l,n)>0$ for $\{l,n\}\in E(j)$, satisfying the conditions on the connectivity graph Laplacian required in Theorem 5.3.4.



(a) (b) (c)

Figure 5.6: Performance of the distributed tracking with consensus algorithm for complete data and noiseless communication case – (a) node consensus estimates $\overline{\mathbf{x}}_n(k, J)$ versus tracking time step, (b) node estimates $\overline{\mathbf{x}}_n(k, j)$ versus consensus iteration number, (c) variance of node estimates $var(k, j)$ versus consensus iteration number.



(a) (b)

Figure 5.7: Steady-state performance of the distributed tracking with consensus algorithm for complete data and noiseless communication case – (a) prediction covariance matrix $\hat{P}_n(k|k-1)$, (b) Kalman gain $K_n(k)$.

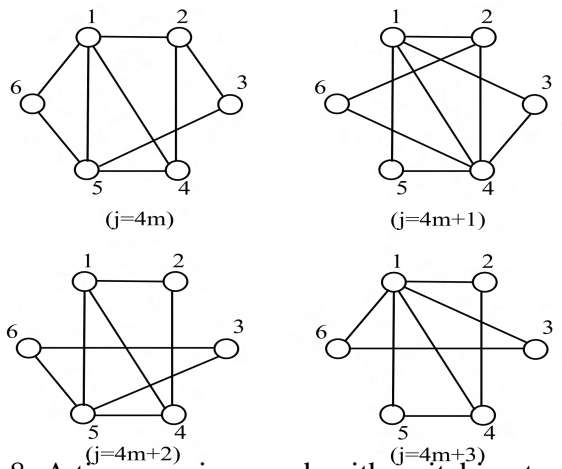
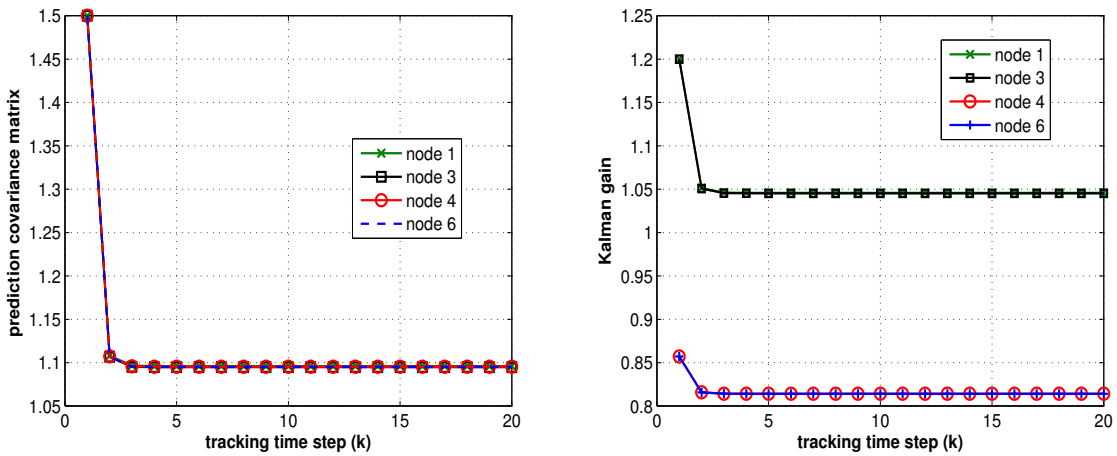


Figure 5.8: A time-varying graph with switching topologies.

Figure 5.9 shows the prediction covariance matrix $\hat{P}_n(k|k-1)$ and Kalman gain $K_n(k)$ of the filter in (5.20), respectively. It can be seen that as the Kalman filter reaches the steady-state, and both the prediction covariance matrix and the Kalman gain converge, as asserted by Theorem 5.3.7. Note that the limit of the Kalman gain is different for different nodes in Fig. 5.9 because the observation matrix H_n is different for different nodes.

5.5 Conclusions

In this chapter, we considered the problem of distributed tracking with consensus on a time-varying graph with incomplete data and noisy communication links. We developed a framework consisting of tracking and consensus updates to handle the issues of time-varying network topology and incomplete data. We discussed the conditions for achieving consensus, quantified the convergence rate and analyzed the steady-state performance when applicable. Our simulation results showed that the proposed distributed tracking with consensus algorithm improves the estimation quality at each node and its performance is close to that of the distributed local Kalman filtering with centralized fusion. The proposed algorithm shows advantages of fully distributed implementation, robustness and scalability, which is preferable in practical application.



(a) (b)

Figure 5.9: Steady-state performance of the distributed tracking with consensus algorithm for incomplete data and noiseless communication case – (a) prediction covariance matrix $\hat{P}_n(k|k-1)$, (b) Kalman gain $K_n(k)$.

Algorithm 1 Distributed Tracking with Consensus Algorithm

Initialize: $\mathbf{x}(0)$, F, H_n, Q,R_n

while new data exists **do**

Kalman filtering in tracking process:

$$\hat{\mathbf{x}}_n(k|k-1) = F\overline{\mathbf{x}}_n(k-1|k-1)$$
$$\hat{P}_n(k|k-1) = F\overline{P}_n(k-1|k-1)F^T + Q$$
$$K_n(k) = \hat{P}_n(k|k-1)H_n^T\left(H_n\hat{P}_n(k|k-1)H_n^T + R_n\right)^{-1}$$
$$\hat{\mathbf{x}}_n(k|k) = \hat{\mathbf{x}}_n(k|k-1) + K_n(k)\Big(\mathbf{y}_n(k) - H_n\hat{\mathbf{x}}_n(k|k-1)\Big)$$
$$\hat{P}_n(k|k) = \Big(I - K_n(k)H_n\Big)\hat{P}_n(k|k-1)$$

update the initial state of consensus process:

$$\overline{\mathbf{x}}_n(k,0) = \hat{\mathbf{x}}_n(k|k)$$
$$\overline{P}(k,0) = \hat{P}_1(k|k) \oplus \hat{P}_2(k|k) \oplus \cdots \oplus \hat{P}_N(k|k)$$
$$j \leftarrow 0$$

while $j \leq J-1$ **do**

$$\overline{\mathbf{x}}_n(k,j+1) = \overline{\mathbf{x}}_n(k,j) + \gamma_n(j)\sum_{l=1}^{N}\mathtt{A}_{n,l}(j)\Big(\mathbf{z}_{n,l}(k,j) - \mathbf{z}_{n,n}(k,j)\Big)$$
$$j \leftarrow j+1$$

end while

$$\overline{\mathbf{x}}_n(k|k) = \overline{\mathbf{x}}_n(k,J)$$
$$\overline{P}_n(k|k) = \overline{P}_n(k,J)$$
$$k \leftarrow k+1$$

end while

Chapter 6

Conclusions and Future Work

In this dissertation we considered the application of information theory and distributed control to different types of networked multi-agent systems, such as wireless sensor networks, networked leader-follower systems, smart grid with DGs and loads, automated highway systems with inter-vehicle communication.

In Chapter 2 we considered tracking in leader-follower systems under communication constraints, where the system components are distributed and connected over communication links with finite data rates. By using information theory and control theory, we provided lower bounds on the channel rate of each communication link as necessary conditions for tracking in such a leader-follower system. We presented limitations in each feedback and forward channel and both channels as a cascade link. Although we only applied simple system models, the results also provide fundamental limitations in terms of information quantities on communication links which can have important roles on control design in leader-follower systems and it should be taken into account for designing new control system with communication constraints. Another limitation of our leader-follower system model is that it only includes a leader and a follower. Our future work is to extend the leader-follower system to more general framework in which multiple leaders and followers are interconnected as a network with more general graph topologies.

In Chapter 3 we proposed a hierarchical decision-making and control architecture for smart grid in which distributed customers equipped with RDG interact and trade energy in the grid. Within this framework, we proposed a distributed networked control strategy with prediction to solve the demand-side management problem encountered within a microgrid with time delay. The approaches we proposed here are distributed, fully scalable and easy to implement, which provides nearly perfect performance with the cost of communication. But in this approach we do not consider economic factors such as pricing information, storage cost and utilities of the adjustable loads, etc, which could be an extension of this work. On the other hand, the method we proposed demands information exchange among users and a communication infrastructure is required with communication cost. Possible further work involves relaxing the assumption of information exchange among users and designing a suboptimal but more practical control algorithm to reduce information exchange.

In Chapter 4 we introduced a leader-following consensus algorithm with communication network in vehicle platoons to guarantee string stability. We showed the sufficient conditions on the eigenvalues of communication graph Laplacian matrix and the control parameters in each vehicle for string stability by using the leader-following consensus algorithm. From simulation results, our proposed approach shows advantage in term of distributed processing, scalability and simpler coordination requirements compared to the previous proposed control methods. However, in our simulation, we only considered random graph topology model. In practical, due to distance separation, more general communication graph model should be included, which is a possible extension.

In Chapter 5, we considered the problem of distributed tracking with consensus on a time-varying graph with incomplete data and noisy communication links. We developed a framework consisting of tracking and consensus updates to handle the issues of time-varying network topology and incomplete data. We discussed the conditions for achieving consensus, quantified the convergence rate and analyzed the steady-state performance when applicable. Our simulation results showed that the proposed distributed tracking with consensus algorithm improves the estimation quality at each node and its performance is close to that of the distributed local Kalman filtering

with centralized fusion. The proposed algorithm shows advantages of fully distributed implementation, robustness and scalability, which is preferable in practical application. The drawback of our approach is the communication cost during consensus process, in which each node needs to exchange its local data with every other node in the network. An possible extension is to shorten the iteration of consensus process or compress the local data that is to be transmitted in order to reduce communication cost.

Appendices

Appendix A

Proof of Lemma 4.3.3:

Proof. The matrix M is non-singular, since $\det(M) \neq 0$. Define $R = M^T M$. Since M is a square matrix and nonsingular, R is positive definite [156]. From the relation between the singular value decomposition (SVD) of M and the eigenvalue-decomposition of $M^T M$ [156], we have $\sigma_i(M) = \sqrt{\lambda_i(M^T M)} = \sqrt{\lambda_i(R)}$. From Gersgorin circle theorem [156], we have $|\lambda_i(R) - r_{i,i}| \leq G(R)$, where $G(R) = \sum_{j \neq i}^{n} |r_{i,j}|$ and $r_{i,j}$ is the (i,j)th element of matrix R. Hence, $0 \leq \lambda_i(R) \leq 4$. Then, we have $0 \leq \sigma_i(M) \leq 2$. □

Appendix B

Proof of Lemma 5.3.3:

Proof. Since $\lambda_2(\overline{L}) > 0$ and $p(l,n) > 0$ for $\{l,n\} \in E(j)$, the undirected time-varying connectivity graph $G(j)$ is connected on average with non-zero link probability. For j large enough, each node will receive the information from one another and generate its updated local estimates. For a fixed k, let $J_k = \inf\{j|(S_k^j)^c = \emptyset, j \geq 0\}$. Then, we have the effective network graph is the same as connectivity graph $\tilde{G}(j) = G(j)$, $\mathtt{L}(j) = L(j)$ and $\Gamma(j) = \gamma(j)I_N$ for $j \geq J_k$.

Since $\mathbb{P}\big[\overline{\mathbf{X}}(j)|\overline{\mathbf{X}}(j-1), \cdots, \overline{\mathbf{X}}(0)\big] = \mathbb{P}\big[\overline{\mathbf{X}}(j)|\overline{\mathbf{X}}(j-1)\big]$, the process $\big\{\overline{\mathbf{X}}(j)\big\}_{j\geq 0}$ is Markov. Define $V(j,\overline{\mathbf{X}}) = \overline{\mathbf{X}}^T(\overline{L} \otimes I_M)\overline{\mathbf{X}}$. Since we assume the graph is undirected and connected on average, $\overline{L}$ is positive semi-definite. Then, the potential function $V(j,\overline{\mathbf{X}})$ is nonnegative. Since $\overline{\mathbf{X}} \in \mathcal{C}$ is an eigenvector of $\overline{L} \otimes I_M$ with zero eigenvalue, $V(j,\overline{\mathbf{X}}) \equiv 0, \overline{\mathbf{X}} \in \mathcal{C}, \lim_{\overline{\mathbf{X}}\to\mathcal{C}} \sup_{j\geq 0} V(j,\overline{\mathbf{X}}) = 0$. From Courant-Fisher Theorem [157, 168], for $\mathbf{Z} \in \mathbb{R}^{NM}$ and $\mathbf{Z} \perp \mathcal{C}$, we have

$$\mathbf{Z}^T(\overline{L} \otimes I_M)\mathbf{Z} \geq \lambda_2(\overline{L} \otimes I_M)\mathbf{Z}^T\mathbf{Z}. \tag{B.1}$$

From Lemma 5.3.1 and the complement of the ϵ-neighborhood of a set in (5.10), we have $\overline{\mathbf{X}} \in V_\epsilon(\mathcal{C}) \Rightarrow \|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\| \geq \epsilon$. Then for $\overline{\mathbf{X}} \in V_\epsilon(\mathcal{C})$, from (B.1) and the properties of Kronecker product

Appendix B. Proof of Lemma 5.3.3:

and eigenvalues, we will have

$$\begin{aligned} V(j,\overline{\mathbf{X}}) =&\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)\overline{\mathbf{X}} = \overline{\mathbf{X}}^T_{\mathcal{C}^\perp}(\overline{L}\otimes I_M)\overline{\mathbf{X}}_{\mathcal{C}^\perp} + \overline{\mathbf{X}}^T_{\mathcal{C}}(\overline{L}\otimes I_M)\overline{\mathbf{X}}_{\mathcal{C}} \geq \lambda_2(\overline{L}\otimes I_M)\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2, \\ =&\lambda_2(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2 \geq \lambda_2(\overline{L})\epsilon^2. \end{aligned} \tag{B.2}$$

Since $\lambda_2(\overline{L}) > 0$, we get $\inf_{j\geq 0,\overline{\mathbf{X}}\in V_\epsilon(\mathcal{C})} V(j,\overline{\mathbf{X}}) \geq \lambda_2(\overline{L})\epsilon^2 > 0$. Consider the generating operator $\mathcal{L}$ and (5.11). Using the fact that $\mathtt{L}(j) = L(j)$ for $j \geq J_k$, we obtain

$$\begin{aligned} \mathcal{L}V(j,\overline{\mathbf{X}}) =&E\left[\overline{\mathbf{X}}(j+1)^T(\overline{L}\otimes I_M)\overline{\mathbf{X}}(j+1)|\overline{\mathbf{X}}(j)=\overline{\mathbf{X}}\right] - \overline{\mathbf{X}}^T(\overline{L}\otimes I_M)\overline{\mathbf{X}}, \\ =&E\Big[\left[\overline{\mathbf{X}} - \gamma(j)\big(L(j)\otimes I_M\big)\overline{\mathbf{X}} - \gamma(j)\overline{\Phi}(j)\right]^T(\overline{L}\otimes I_M) \\ &\times\left[\overline{\mathbf{X}} - \gamma(j)\big(L(j)\otimes I_M\big)\overline{\mathbf{X}} - \gamma(j)\overline{\Phi}(j)\right]\Big] - \overline{\mathbf{X}}^T(\overline{L}\otimes I_M)\overline{\mathbf{X}} \quad \text{for } j\geq J_k. \end{aligned}$$

From (5.6), we have $E\big[\|\overline{\Phi}(j)\|^2\big] \leq \eta$. By using the independence of $L(j)$ and $\overline{\Phi}(j)$ with respect to $\overline{\mathbf{X}}(j)$ and $\overline{\mathbf{X}}^T\overline{L}\overline{\mathbf{X}} \leq \lambda_N(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2$ [158], after some work we have that

$$\begin{aligned} &\mathcal{L}V(j,\overline{\mathbf{X}}) = -2\gamma(j)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^2\overline{\mathbf{X}} + \gamma^2(j)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^3\overline{\mathbf{X}} \\ &+ E\Big[\gamma^2(j)\big[\big(\tilde{L}(j)\otimes I_M\big)\overline{\mathbf{X}}\big]^T(\overline{L}\otimes I_M)\big(\tilde{L}(j)\otimes I_M\big)\overline{\mathbf{X}}\Big] + E\big[\gamma^2(j)\overline{\Phi}(j)^T(\overline{L}\otimes I_M)\overline{\Phi}(j)\big], \\ &\leq -2\gamma(j)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^2\overline{\mathbf{X}} + \gamma^2(j)\left[\lambda_N^3(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2\right. \\ &+ \lambda_N(\overline{L})E\big[\|\big(\tilde{L}(j)\otimes I_M\big)\overline{\mathbf{X}}\|^2\big] + \lambda_N(\overline{L})E\big[\|\overline{\Phi}(j)\|^2\big]\Big], \\ &\leq -2\gamma(j)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^2\overline{\mathbf{X}} + \gamma^2(j)\left[\lambda_N^3(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2 + 4N^2\lambda_N(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2 + \lambda_N(\overline{L})\eta\right]. \end{aligned}$$

The last step follows from the fact that all the eigenvalues of $\tilde{L}(j)$ are less than $2N$ in absolute value, by the Gershgorin circle theorem. Using the fact that $\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)\overline{\mathbf{X}} \geq \lambda_2(\overline{L})\|\overline{\mathbf{X}}_{\mathcal{C}^\perp}\|^2$ from (B.2), we have

$$\begin{aligned} \mathcal{L}V(j,\overline{\mathbf{X}}) \leq& -2\gamma(j)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^2\overline{\mathbf{X}} + \gamma^2(j)\left[\lambda_N(\overline{L})\eta + \left(\frac{\lambda_N^3(\overline{L})}{\lambda_2(\overline{L})} + \frac{4N^2\lambda_N(\overline{L})}{\lambda_2(\overline{L})}\right)\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)\overline{\mathbf{X}}\right], \\ \leq& -2\gamma(j)\varphi(j,\overline{\mathbf{X}}) + g(j)\big[1+V(j,\overline{\mathbf{X}})\big] \quad \text{for } j\geq J_k, \end{aligned}$$

where $\varphi(j,\overline{\mathbf{X}}) = 2\overline{\mathbf{X}}^T(\overline{L}\otimes I_M)^2\overline{\mathbf{X}}$, $g(j) = \gamma^2(j)\max\left(\lambda_N(\overline{L})\eta, \frac{\lambda_N^3(\overline{L})}{\lambda_2(\overline{L})} + \frac{4N^2\lambda_N(\overline{L})}{\lambda_2(\overline{L})}\right)$. Then, the theorem follows by using Lemma 5.3.2. □

Appendix C

Proof of Theorem 5.3.7:

Proof. Step 1: Bound on the error covariance

From (5.1), it can be easily shown that the controllability matrix has full rank and the system is controllable. Since (F, H_n) is detectable, $\exists K_n'$ such that $(F - K_n' H_n)$ are stable. Consider the sub-optimal filter

$$\hat{x}_n(k+1|k) = F\hat{x}_n(k|k-1) + K_n'\big(y_n(k) - H_n\hat{x}_n(k|k-1)\big),$$

Since consensus is reached in consensus update part, $\hat{x}_n(k|k-1) = \hat{x}_l(k|k-1) = \hat{x}(k|k-1)$ for $1 \leq n, l \leq N$. Then,

$$\hat{x}(k+1|k) = \frac{1}{N}\left(F\sum_{n=1}^{m}\beta_n - \sum_{n=1}^{m}K_n' H_n\beta_n\right)\hat{x}(k|k-1) + \frac{1}{N}\sum_{n=1}^{m}K_n' y_n(k)\beta_n.$$

It is easily verified that

$$\begin{aligned}\tilde{x}(k+1|k) &= x\big(k+1\big) - \hat{x}(k+1|k),\\ &= \left(F - \frac{1}{N}\sum_{n=1}^{m}K_n' H_n\beta_n\right)\tilde{x}(k|k-1) - \frac{1}{N}\sum_{n=1}^{m}K_n' v_n(k)\beta_n + w(k),\end{aligned}$$

where the last step follows from the fact that the estimate is unbiased and $\frac{1}{N}\sum_{n=1}^{m}\beta_n = 1$. Since $(F - K_n' H_n)$ is stable, $\big(F - \frac{1}{N}\sum_{n=1}^{m} K_n' H_n \beta_n\big)$ is also stable. It follows that the covariance matrix $\Pi(k) = \text{Cov}\big[\tilde{x}(k|k-1)\big]$ is bounded, where $\text{Cov}(x)$ denotes the covariance matrix of x. However, the filter above is sub-optimal, so that $P(k|k-1) \leq \Pi(k)$.

Step 2: Monotonicity of the error covariance

Recall that the mapping $f : \hat{P}_n(k|k-1) \to \hat{P}_n(k+1|k)$ as $\hat{P}_n(k+1|k) = \min_{K_n} g\big(\hat{P}_n(k|k-1), K_n\big)$, where

$$g(\hat{P}_n, K_n) = (F - K_n H_n)\hat{P}_n(F - K_n H_n)^T + K_n R_n K_n^T + Q.$$

Thus, if $\hat{P}_n(k|k-1) \geq \hat{P}_n'(k|k-1)$,

$$\begin{aligned}\hat{P}_n(k+1|k) &= \min_{K_n} g\big(\hat{P}_n(k|k-1), K_n\big) = g\big(\hat{P}_n(k|k-1), K_n^*\big) \geq g\big(\hat{P}_n'(k|k-1), K_n^*\big), \\ &\geq \min_{K_n} g\big(\hat{P}_n'(k|k-1), K_n\big) = \hat{P}_n'(k+1|k).\end{aligned}$$

Therefore, the mapping f from $\hat{P}_n(k|k-1)$ to $\hat{P}_n(k+1|k)$ is monotonic. Because $\hat{P}(k+1|k) = \frac{1}{N^2}\sum_{n=1}^{m}\hat{P}_n(k+1|k)\beta_n^2$, the mapping $\hat{f} : \hat{P}(k|k-1) \to \hat{P}(k+1|k)$ is also monotonic.

Step 3: Use of zero initial covariance

Suppose $\hat{P}(0|-1) = 0$. Then $\hat{P}(1|0) \geq \hat{P}(0|-1) = 0$. But from Step 2 it follows that $\hat{P}(k+1|k) \geq \hat{P}(k|k-1)$, for $k \geq 0$. Since $\Big\{\hat{P}(k|k-1)\Big\}$ is bounded by Step 1, then $\hat{P}(k|k-1) \to P$ for some $P \geq 0$. Obviously, P must be a stationary point of the covariance update equation, hence solves the DARE.

Step 4: Asymptotic stability of the filter

With $\overline{K}_n$ the stationary gain corresponding to P, the DARE is

$$P = \left(F - \frac{1}{N}\sum_{n=1}^{m}\overline{K}_n H_n \beta_n\right) P \left(F - \frac{1}{N}\sum_{n=1}^{m}\overline{K}_n H_n \beta_n\right)^T + \frac{1}{N^2}\sum_{n=1}^{m}\overline{K}_n R_n \overline{K}_n^T \beta_n^2 + GG^T,$$

Appendix C. Proof of Theorem 5.3.7:

where $GG^T = Q$. Let ν be a left eigenvector of $\left(F - \frac{1}{N}\sum_{n=1}^{m} \overline{K}_n H_n \beta_n\right)$ with eigenvalue λ. Then

$$\left(\nu P \nu^T\right) = |\lambda|^2 \left(\nu P \nu^T\right) + \frac{1}{N^2} \sum_{n=1}^{m} \nu \overline{K}_n R_n \overline{K}_n^T \nu^T \beta_n^2 + \nu G G^T \nu^T. \tag{C.1}$$

Since R_n and Q are positive semidefinite, it implies that $|\lambda| \leq 1$. It only remains to show that $|\lambda| = 1$ is impossible. If $|\lambda| = 1$, we have from (C.1) and the definition of ν:

$$\nu \left(F - \frac{1}{N} \sum_{n=1}^{m} \overline{K}_n H_n \beta_n \right) = \lambda \nu, \qquad \nu \overline{K}_n = 0, \qquad \text{and} \qquad \nu G = 0,$$

which gives that $\nu[\lambda I - F, G] = 0$. This contradicts the assumption that (F, G) is stabilizable.

Step 5: Nonzero initial covariances

Suppose we use the stationary suboptimal filter $K'_n \equiv \overline{K}_n$ to obtain the estimate $\hat{x}(k|k-1)$. We show that its error covariance converges to P. Defining $\tilde{x}(k|k-1) \triangleq x(k) - \hat{x}(k|k-1)$ we obtain

$$\tilde{x}(k|k-1) = \left(F - \frac{1}{N} \sum_{n=1}^{m} \overline{K}_n H_n \beta_n \right) \tilde{x}(k|k-1) - \frac{1}{N} \sum_{n=1}^{m} \overline{K}_n v_n(k) \beta_n + w(k).$$

Since $(F - \frac{1}{N}\sum_{n=1}^{m} \overline{K}_n H_n \beta_n)$ is stable with eigenvalue $|\lambda| < 1$, it follows from above results on stationary behavior that $\Pi(k) \equiv \text{Cov}[\tilde{x}(k|k-1)] \to \tilde{P} \geq 0$, where $\tilde{P}$ is the unique non-negative solution of the Lyapunov equation:

$$\tilde{P} = \left(F - \frac{1}{N} \sum_{n=1}^{m} \overline{K}_n H_n \beta_n \right) \tilde{P} \left(F - \frac{1}{N} \sum_{n=1}^{m} \overline{K}_n H_n \beta_n \right)^T + \frac{1}{N^2} \sum_{n=1}^{m} \overline{K}_n R_n \overline{K}_n^T \beta_n^2 + Q.$$

However, substituting $\overline{K}_n$ this is just the DARE which is satisfied by P, hence $\tilde{P} = P$. Now $\hat{x}(k|k-1)$ is sub-optimal so that $P(k|k-1) \leq \Pi(k) \to \tilde{P}$. On the other hand, by monotonicity of mapping $\hat{f} : \hat{P}(k|k-1) \to \hat{P}(k+1|k)$, it follows that $P(k|k-1) \geq P^0(k|k-1) \to P$, where $P^0(k|k-1)$ is the covariance for $P(0|-1) = 0$. Hence, $P(k|k-1) \to P$. □

References

[1] K. Manickavasagam, M. Nithya, K. Priya, J. Shruthi, S. Krishnan, S. Misra, and S. Manikandan, "Control of distributed generator and smart grid using multi-agent system," in *Proceedings of 2011 1st International Conference on Electrical Energy Systems (ICEES)*, pp. 212-217, Jan. 2011.

[2] A. L. Dimeas and N. D. Hatziargyriou, "Operation of a multi-agent system for microgrid control," *IEEE Transactions Power System*, vol. 20, no. 3, pp. 1447-1455, Aug. 2005.

[3] S. D. J. McArthur, E. M. Davidson, V. M. Catterson, A. L. Dimeas, N. D. Hatziargyriou, F. Ponci, T. Funabashi, "Multi-agent systems for power engineering applications – Part I: Concepts, approaches, and technical challenges," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1743-1752, Nov. 2007.

[4] S. D. J. McArthur, E. M Davidson, V. M Catterson, A. L. Dimeas, N. D. Hatziargyriou, F. Ponci, T. Funabashi, "Multi-agent systems for power engineering applications– Part II: Technologies, standards, and tools for building multi-agent systems," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1753-1759, Nov. 2007.

[5] L.Peppard, "String stability of relative-motion pid vehicle control systems," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 579–581, Oct. 1974.

[6] D. Swaroop and J. K. Hedrick, "String Stability of Interconnected Systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp.349–357, 1996.

[7] D. Swaroop and J. K. Hedrick, "Constant spacing strategies for platooning in automated highway systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 121, no. 3, pp. 462–470, 1999.

[8] P. Seiler, A. Pant, and K. Hedrick, "Disturbance propagation in vehicle strings," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1835-1841, Oct. 2004.

[9] P. A. Cook, "Stable control of vehicle convoys for safety and comfort, *IEEE Transactions on Automatic Control*, vol. 52, no. 3, pp. 526-531, Mar. 2007.

[10] J. C. Hill and J. K. Archibald and W. C. Stirling and R. L. Frost, "A multi-agent system architecture for distributed air traffic control," in *Proceedings of AIAA Guidance, Navigation and Control Conference*, pp. 2005-6049, 2005.

[11] K. Tumer, A. K. Agogino, "Improving air traffic management with a learning multi-agent system," *IEEE Intelligent Systems Magazine*, vol. 24, no.1, pp. 18-21, 2009.

[12] K. Hwang, S. Tan, M. Hsiao, C. Wu, "Cooperative multiagent congestion control for high-speed networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 2, pp. 255-268, April 2005.

[13] R. Che, *Dynamical congestion control strategies for a network of multi-agent systems subject to differentiated services traffic*. Ph.D. dissertation, Concordia University, 2011.

[14] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *System Control Letter*, vol. 53, pp. 65-78, 2004.

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

[16] M.l Vinyals, J. A. Rodrguez-Aguilar, J. Cerquides, "A survey on sensor networks from a multiagent perspective," *Computer Journal*, vol. 54, no. 3, pp. 455-470, 2011.

[17] C. V. Goldman and S. Zilberstein, "Decentralized control of control of cooperative systems: Categorization and complexity analysis", *Journal of Artificial Intelligence Research*, vol. 22, pp. 143-174, 2004.

[18] R. M. Murray, K. J. Astrom, S. P. Boyd, R. W. Brockett, G. Stein, "Future directions in control in an information-rich world," *IEEE Control Systems Magazine*, vol. 23, no. 2, pp. 20-33, Apr 2003.

[19] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138-162, Jan. 2007.

[20] T. C. Yang, "Networked control system: a brief survey," *IEE Proceedings - Control Theory and Applications*, vol. 153, no. 4, pp. 403-412, July 2006.

[21] X. M. Tang, J. S. Yu, *Bio-Inspired Computational Intelligence and Applications*. Berlin Heidelbery: Springer, 2007.

[22] R. A. Gupta, M. Chow, "Networked control system: Overview and research trends," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 7, pp. 2527-2535, July 2010.

[23] S. Tatikonda, S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056-1068, 2004.

[24] S. Tatikonda, S. Mitter. "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196-1201, 2004.

[25] S. Tatikonda, A. Sahai, S. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, 1549-1561, 2004.

[26] G. N. Nair, R. J. Evans. "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control Optimization*, vol. 43, no. 2, pp. 413-436, 2004.

[27] G. N. Nair, F. Fagnani, S. Zampieri, R. J. Evans, "Feedback control under data rate constraints: an overview", *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108-137, Jan. 2007.

[28] P. Minero, M. Franceschetti, S. Dey and G.N. Nair, "Data rate theorem for stabilization over time-varying feedback channels", *IEEE Trans. Automatic Control*, vol. 54, no. 2, pp. 243-255, Feb. 2009.

[29] N. C. Martins, M. A Dahleh, "Feedback control in the presence of noisy channels: 'Bode-Like' fundamental limitations of performance", *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp.1604-1615, 2004.

[30] N. C. Martins, M. A. Dahleh, N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 438-447, Mar. 2006.

[31] N. C. Martins, M. A Dahleh, J.C. Doyle, "Fundamental limitations of disturbance attenuation in the presence of side information," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 56-66, 2007.

[32] V. Gupta and N. Martins, "On stability in the presence of analog erasure channels between controller and actuator," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 175-179, Jan. 2010.

[33] L. Shi, R. M. Murray, "Towards a packet-based control theory - Part 1: stabilization over a packet-based network," in *Proceedings of the American Control Conference*, pp. 1251-1256, June 2005.

[34] L. Shi, R. M. Murray, "Towards a packet-based control theory - Part 2: rate issues," in *Proceedings of the American Control Conference*, pp.3482-3487, June 2006.

[35] A. Sahai, S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link - Part I: scalar systems," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3369-3395, Aug. 2006.

[36] A. J. Rojas, J. H. Braslavsky, R. H. Middleton, "Output feedback stabilisation over bandwidth limited, signal to noise ratio constrained communication channels," in *Proceedings of the American Control Conference*, pp. 6, June 2006.

[37] A. J. Rojas, R. H. Middleton, J. S. Freudenberg, J. H. Braslavsky, "Input disturbance rejection in channel signal-to-noise ratio constrained feedback control," in *Proceedings of the American Control Conference*, pp. 3100-3105, June 2008.

[38] A. J. Rojas, "Signal-to-noise ratio fundamental limitations in continuous-time linear output feedback control," *IEEE Transactions on Automatic Control*, vol. 54, no. 8, pp. 1902-1907, Aug. 2009.

[39] A. J. Rojas, "Signal-to-noise ratio constrained feedback control over fading Channels," in *Proceedings of the 18th IFAC World Congress*, pp. 8799-8804, Milano, Italy, Aug. 2011.

[40] A. J. Rojas, "Signal-to-noise ratio limitations in feedback control," *IEEE Latin America Transactions (Revista IEEE America Latina)*, vol. 9, no. 5, pp. 690-699, Sept. 2011.

[41] H. Zhang, Y. Sun, "Bode integrals and laws of variety in linear control systems," in *Proceedings of the IEEE American Control Conference* pp. 66-70, Piscataway, CA, June 2003.

[42] H. Zhang and Y. Sun, "Information theoretic interpretations for H_∞ entropy, in *Proceedings of the 16th IFAC World Congress*, Prague, July 2005.

[43] H. Zhang and Y. Sun, "Epsilon-entropy and H_∞ entropy in continuous time systems, in *Proceedings of the 17th IFAC World Congress*, pp. 8021-8026, Souel, Korea, July, 2008.

[44] H. Zhang and Y. Sun, "Variety of linear continuous time systems, in *Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, Shanghai, P.R. China, pp. 6426-6431, Dec. 2009.

[45] H. Kwakernaak, and R. Sivan, *Linear Optimal Control Systems*. Wiley-Interscience, 1972.

[46] P. Dorato, C. T. Abdallah, and V. Cerone, *Linear Quadratic Control: An Introduction*. Prentice Hal, 1995.

[47] V. Gupta, N.C. Martins, and J.S. Baras, "Optimal output feedback control using two remote sensors over erasure channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp.1463-1476, July 2009.

[48] V. Gupta and N. C. Martins, "Optimal tracking control across erasure communication links, in the presence of preview," *International Journal for Robust and Nonlinear Control*, vol. 19, no. 16, pp. 1837-1850, Nov. 2009.

[49] A. Sahai, ""Any-time" capacity and a separation theorem for tracking unstable processes," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 500, 2000.

[50] A. Sahai, *Anytime Information Theory*. Ph.D. Dissertation, MIT 2001.

[51] S. Park, A. Sahai, "Network coding meets decentralized control: Capacity-stabilizabililty equivalence," in *Proceedings of the 2001 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pp. 4817-4822, Dec. 2011.

[52] H. Zhang Y. Sun, "Information theoretic limit and bound of disturbance rejection in LTI systems: Shannon entropy and H_∞ entropy," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 2, pp. 1378-383, Oct. 2003.

[53] H. Zhang, Y. Sun, "Directed information and mutual information in linear feedback tracking systems," in *Proceedings of the Sixth World Congress onIntelligent Control and Automation (WCICA 2006)*, vol. 1, pp. 723-727, 2006.

[54] H. Zhang and Y. Sun, "On information transmission in linear feedback tracking systems," *Asia-Pacific Journal of Chemical Engineering*, vol. 3, no. 6, pp. 630-637, Nov. 2008.

[55] N. van de Wouw, P. Naghshtabrizi, M. Cloosterman, J. P. Hespanha, "Tracking control for networked control systems," in *Proceedings of the 2007 46th IEEE Conference on Decision and Control*, pp. 4441-4446, Dec. 2007.

[56] I. Lopez, C. T. Abdallah, S. K. Jayaweera, H. Tanner, "Conditions for tracking in networked control systems," in *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 3626-3632, Dec. 2008.

[57] C. I. L. Hurtado, C. T. Abdallah, S. K. Jayaweera. "Limitations in tracking systems," *Journal of Control Theory and Applications*, vol. 8, no. 3, pp. 351-358, 2010.

[58] C. I. L. Hurtado, *Limitations in tracking systems*. Ph.D dissertation, University of New Mexico, 2010.

[59] S. M. Amin and B.F. Wollenberg, "Toward a smart grid," *IEEE Power & Energy Magazine*, vol. 3, no. 5, pp. 34-41, Sept.-Oct. 2005.

[60] U.S. Dept. of Energy, "The smart grid: an introduction," Tech. Report, 2008.

[61] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids, *IEEE Power & Energy Magazine*, vol. 5, no. 4, pp. 78-94, July-Aug. 2007.

[62] X. Fang, S. Misra, G. Xue and D. Yang, "Smart grid-The new and improved power grid: A survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 994-980, Fourth Quarter 2012.

[63] N. D. Hatziargyriou, "Microgrids, *IEEE Power & Energy Magazine*, vol. 6, no. 3, pp. 26-29, May-June 2008.

[64] V. Bakker, M. Bosman, A. Molderink, J. Hurink, and G. Smit. "Demand-side load management using a three step optimization methodology," in *Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications*, pp. 431-436, Gaithersburg, Maryland, Oct. 2010.

[65] D. ONeill, M. Levorato, A. Goldsmith, and U. Mitra, "Residential demand response using reinforcement learning," in *Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications*, pp. 409-414, Gaithersburg, Maryland, Oct. 2010.

[66] X. Guan, Z. Xu, and Q.S. Jia, "Energy-efficient buildings facilitated by microgrid, *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 243-252, Dec. 2010.

[67] A. Conejo, J. Morales, and L. Baringo, "Real-time demand response model," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 236-242, Dec. 2010.

[68] D. Li, S. K. Jayaweera, "Distributed Smart-home Decision-making in a Hierarchical Interactive Smart Grid Architecture," *IEEE Transactions on Parallel and Distributed Systems*, to be appeared, 2014.

[69] A. H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320-331, Dec. 2010.

[70] N. Gatsis, G. Giannakis, "Cooperative multi-residence demand response scheduling," in *Proceedings of the 2011 45th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, Mar. 2011.

[71] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Message passing for dynamic network energy management", 2012 [Online]. Available: http:// arxiv.org/pdf/1204.1106.pdf.

[72] D. Swaroop, *String stability of interconnected systems: An application to platooning In automated highway systems.* Ph.D Dissertation, University of California, Berkeley, 1997.

[73] D. YanaKiev, I. Kanellakopoulos, "A simplified framework for string stability analysis in AHS," in *Proceedings of the 13th IFAC World Congress*, pp. 177-182, 1996.

[74] S. Mahal, *Effects of communication delays on string stability in an AHS environment.* Master Thesis, University Of California, Berkeley, 2000.

[75] P. Fernandes and U. Nunes, "Platooning of autonomous vehicles with intervehicle communications in SUMO traffic simulator," in *Proceedings of the 2010 13th International IEEE Annual Conference on Intelligent Transportation Systems*, Madeira Island, Portugal, Sept. 2010.

[76] P. Ioannou,and C. C. Chien, "Autonomous intelligent cruise control", *IEEE Transactions Vehicular Technology*, vol. 42, no. 4, pp. 657-672, 1993.

[77] L. Xiao; S. Darbha, F. Gao, "Stability of string of adaptive cruise control vehicles with parasitic delays and lags," in *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*, pp. 1101-1106, Oct. 2008.

[78] X. Liu, S. Mahal, A. Goldsmith and J. Hedrick, "Effects of communication delays on string stability in vehicle platoons", in *Proceedings of the 2001 IEEE Intelligent Transportation Systems*, pp. 625-630, Oakland, USA, Aug. 2001.

[79] Y. Zhai, *Design of switching strategy for adaptive cruise control under string stability constraints.* Master Thesis, Purdue University, 2011.

[80] W. Levine and M. Athans, "On the optimal error regulation of a string of moving vehicles", *IEEE Transactions Automatic Control*, vol. 11, no. 3, pp. 355-361, July 1966.

[81] M. E. Khatir and E. J. Davison. "Decentralized control of a large platoon of vehicles using non-identical controllers," In *Proceedings of the 2004 American Control Conference*, pp. 2769-2776, Boston, Massachusetts, July 2004.

[82] E. Shaw, *Heterogeneous string stability for vehicle formation control.* Ph.D Dissertation, University Of California, Berkeley, 2005.

[83] E. Shaw and J.K. Hedrick, "Controller design for string stable heterogeneous vehicle strings," In *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 2868-2875, Dec. 2007.

[84] J. A. Rogge, and D. Aeyels, "Vehicle platoons through ring coupling, *IEEE Transactions on Automatic Control*, vol. 53, pp. 1370-1377, July 2008.

[85] S. Klinge, *Stability issues in distributed systems of vehicle platoons.* Master Thesis, Otto-von-Guericke-University, 2008.

[86] S. Klinge, R. H. Middleton, "String stability analysis of homogeneous linear unidirectionally connected systems with nonzero initial conditions," in *Proceedings of the IET Signals and Systems Conference (ISSC 2009)*, Dublin, Irish, pp. 1-6, June 2009.

[87] S. Klinge and R. H. Middleton, "Time headway requirements for string stability of homogeneous linear unidirectionally connected systems," in *Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pp. 1992-1997, Shanghai, P.R. China, Dec. 2009.

[88] R. H. Middleton and J. H. Braslavsky, "String instability in classes of linear time invariant formation control with limited communication range," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1519-1530, 2009.

[89] W. B. Dunbar, D. S. Caveney, "Distributed receding horizon control of vehicle platoons: Stability and string stability," *IEEE Transactions on Automatic Control*, vol. 57, no. 3, pp. 620-633, Mar. 2012.

[90] S. Sheikholeslam, and C. A. Desoer, "Longitudinal Control of a Platoon of Vehicles", in *Proceedings of 1990 American Control Conference*, pp. 291-296, San Diego, CA, 1990.

[91] B. D. O. Anderson, and J. B. Moore, *Optimal Filtering*. New Jersey: Prentice-Hall. Englewood Cliffs, 1979.

[92] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. on Automatic Control*, vol. 24, no. 6, pp. 843-854, Dec. 1979.

[93] Y. Bar-Shalom, and T. E. Fortmann, *Tracking and Data Association*. Boston: Academic Press, 1988.

[94] Y. Bar-Shalom, and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[95] M. Fogel, N. Burkhart, H. Ren, J. Schiff, M. Meng, and K. Goldberg, "Automated tracking of pallets in warehouses: Beacon layout and asymmetric ultrasound observation models," in *Proceedings of the Third IEEE Conference on Automation Science and Engineering (CASE)*, pp. 678-685, Scottsdale, AZ, Sep. 2007.

[96] J. Wilson, V. Bhargava, A. Redfern, and P. Wright, "A wireless sensor network and incident command system for urban firefighting," in *Proceedings of the 4th Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services (MobiQuitous)*, pp. 1-7, Philadephia, PA, 2007.

[97] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proceedings of the 5th ACM/IEEE Conf. on Mobile Comp. and Networking.*, pp. 263-270, Seattle, WA, 1999.

[98] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988-1001, Jun. 2003.

[99] W. Ren, R. W. Beard, and E. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proceedings of American Control Conference*, Portland, OR, pp. 1859-1864, Jun. 2005

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

[101] R. Olfati-Saber, and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520-1533, Sep. 2004.

[102] D. B. Kingston and R. W. Beard, "Discrete-time average-consensus under switching network topologies," in *Proceedings of American Control Conference*, Minneapolis, MN, pp. 3551-3556, Jun. 2006.

[103] Y. Hatano, and M. Mesbahi, "Agreement over random networks," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1867-1872, Nov. 2005.

[104] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192-1203, Jul. 2007.

[105] F. Fagnani, and S. Zampieri, "Average consensus with packet drop communication," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 102-133, 2009.

[106] S. Kar, and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3315-3326, Jul. 2008.

[107] S. Kar, and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355-369, Jan. 2009.

[108] M. Huang, and J. H. Manton, "Coordination and consensus of networked agents with noisy measurement: Stochastic algorithms and asymptotic behavior," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 134-161, 2009.

[109] T. Li, "Asymptotically unbiased average consensus under measurement noises and fixed topologies," in *Proceeding of the 17th IFAC World Congress*, Seoul, Korea, pp. 2867-2873, Jul. 2008.

[110] T. Li, and J. F. Zhang, "Mean square average consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929-1936, 2009.

[111] T. Li, and J. F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2043-2057, Sept. 2010.

[112] M. E. Yildiz and A. Scaglione, "Coding with side information for rate-constrained consensus," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3753-3764, 2008.

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

[114] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, pp. 5492-5498, Dec. 2007.

[115] P. Altiksson, and A. Rantzer, "Distributed Kalman filtering using weighted averaging," *Proceedings of the 17th Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, pp. 2445-2450, July 2006.

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

[117] Y. Hong, J. Hu, and L. Gao "Tracking control for multi-agent consensus with an active leader and variable topology," *Automatica*, vol. 42, no. 7, pp. 1177-1182, July 2006.

[118] C. Mosquera, R. Lopez-Valcarce, and S. K. Jayaweera, "Stepsize sequence design for distributed average consensus," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 169-172, Feb. 2010.

[119] S. K. Jayaweera, "Distributed space-object tracking and scheduling with a satellite-assisted collaborative space surveillance network (SSN)," Final project report: AFRL Summer Faculty Fellowship Program, Jul. 2009.

[120] Y. Cao, W. Ren, and Y. Li, "Distributed discrete-time coordinated tracking with a time-varying reference state and limited communication," *Automatica*, vol. 45, no. 5, pp. 1299-1305, May 2009.

[121] S. K. Jayaweera, Y. Ruan and R. S. Erwin, "Distributed Tracking with Consensus on Noisy Time-varying Graphs with Incomplete Data", in *Proceedings of the 10th International Conference on Signal Processing (ICSP2010)*, Beijing, China, pp. 2584-2587, Oct. 2010.

[122] Y. Ruan, S. K. Jayaweera and C. Mosquera, "Performance Analysis of Distributed Tracking with Consensus on Noisy Time-varying Graphs", in *Proceedings of the 5th Advanced Satellite Multimedia System Conference and the 11th Signal Processing for Space Communications Workshop*, Sardinia, Italy, pp. 389-394, Sept. 2010.

[123] K. Topley, V. Krishnamurthy, and G. Yin, "Consensus-tracking in distributed networks by one-hop averaging," in *Proceedings of IEEE International Conference Acoustics, Speech, Signal Processing*, Taipei, Taiwan, pp. 2065-2068, May 2009.

[124] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Grad. Texts in Math. 207, Springer-Verlag, 2001

[125] B. Liu, T. Chu, L. Wang, G. Xie, "Controllability of a leader-follower dynamic network with switching topology," *IEEE Transactions on Automatic Control*, vol. 53, no. 4, pp. 1009-1013, 2008.

[126] Y. Hong, X. Wang, Z. Jiang, "Multi-agent coordination with general linear models: A distributed output regulation approach," in *Proceedings of the IEEE International Conference on Control and Automation (ICCA)*, Ximen, China, pp. 137-142, 2010.

[127] X. Chen, A. Serrani, H. Ozbay, "Control of leader-follower formations of terrestrial UAVs," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, pp. 498-503, 2003.

[128] T. M. Cover, J. A. Thomas, *Elements of Information Theory 2nd ed.* John Wiley & Sons, Hoboken, 2006.

[129] T. E. Duncan, "On the calculation of mutual information," *SIAM Journal on Applied Mathematics*, vol. 19, no. 1, pp. 215-220, 1970.

[130] L. Paninski, "Estimation of entropy and mutual information," *Neural Computation*, vol. 15, no. 6, pp. 1191-1253, 2003.

[131] A. J. Viterbi, J. K Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hi, 1979.

[132] A. V. Oppenheim, R. W. Schafer, J. R. Buck, *Discrete-Time signal processing 2nd ed.* Prentice Hall, 1999.

[133] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes.* San Francisco: Holden-Day, 1964.

[134] E. Koch and M.A.Piette, "Architecture concepts and technical issues for an open, interoperable automated demand response infrastructure, Technical report, Lawrence Berkeley National Laboratory, 2007.

[135] C. H. Hauser, D.E. Bakken and A. Bose, "A failure to communicate: next generation communication requirements, technologies, and architecture for the electric power grid," *IEEE Power and Energy Magazine*, vol. 3, no. 2, pp. 47-55, Mar. 2005.

[136] T. L. Vandoorn, B. Zwaenepoel, J. D. M. De Kooning, B. Meersman and L. Vandevelde, "Smart microgrids and virtual power plants in a hierarchical control structure," in *Proceedings of 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*, pp. 1-7, Dec. 2011.

[137] M. U. Tariq, S. Grijalva and M. Wolf, "Towards a distributed, service-oriented control infrastructure for smart grid," in *Proceedings of 2011 IEEE/ACM International Conference on Cyber-Physical Systems*, pp. 35-44, Apr. 2011.

[138] D. Li, S. K. Jayaweera, and C. Abdallah, "Uncertainty modeling and stochastic control design for smart grid with distributed renewables," in *Proceedings of 2012 IEEE Green Technologies Conference*, Tulsa, OK, pp. 1-3, Apr. 2012.

[139] M. Alizadeh, X. Li, Z. Wang , A. Scaglione, R. Melton, "Demand-side management in the smart grid: Information processing for the power switch," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp.55-67, Sept. 2012.

[140] J. Barton, D. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 441-448, June 2004.

[141] P. R. Kumar, P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, 1986.

[142] H. K. Khalil , *Nonlinear Systems (3rd Edition)*. Prentice Hall, 2001.

[143] H. Kwakernaak, and R. Sivan, *Linear Optimal Control Systems*. Wiley-Interscience, 1972.

[144] B. van Brunt, *The Calculus of Variations*. Springer, 2004.

[145] P. Park, Y. S. Moon and W. H. Kwon, "A stabilizing output-feedback linear quadratic control for pure input-delayed systems," *International Journal of Control*, vol. 7, no. 5, pp. 385-391, 1999.

[146] P. Seiler, A. Pant, and K. Hedrick, "Disturbance propagation in vehicle strings," *IEEE Transactions Automatic Control*, vol. 49, no. 10, pp. 1835-1841, 2004.

[147] Y. Ruan, S. K. Jayaweera and R.S. Erwin, "Distributed tracking-with-consensus on noisy time-varying graphs with incomplete data, *EURASIP Journal on Advances in Signal Processing*, no. 1, pp. 1-21, Nov. 2011.

[148] Z. Meng, W. Ren, Y. Cao and Z. You , "Leaderless and leader-following consensus with communication and input delays under a directed network topology," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 1, pp. 75-88, Feb. 2011.

[149] W. Ni and D. Cheng, "Leader-following consensus of multi-agent systems under fixed and switching topologies," *Systems & Control Letters*, vol. 59, no. 3-4, pp. 209-217, March-April 2010.

[150] R. H. Middleton and J. H. Braslavsky, "String instability in classes of linear time invariant formation control with limited communication range," *IEEE Transactions Automatic Control*, vol. 55, no. 7, pp. 1519-1530, 2010.

[151] J. Hedrick, C. McMahon, and D. Swaroop, "Vehicle modeling and control for automated highway systems," Technical report, University of California, Berkeley, 1993.

[152] P. Dorato, C. T. Abdallah, and V. Cerone, *Linear Quadratic Control: An Introduction*. Prentice Hal, 1995.

[153] H. Knobloch, D. Flockerzi, A. Isidori, *Topics in Control Theory*. Springer, 1993.

[154] P. Barooah and J.P. Hespanha, "Error amplification and disturbance propagation in vehicle strings with decentralized linear control," in *Proceedings of the 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference*, Seville, Spain, pp. 4964-4969, Dec. 2005.

[155] I. Lestas and G. Vinnicombe, "Scalability in heterogeneous vehicle platoons," in *Proceedings of American Control Conference 2007*, New York, NY, pp. 4678-4683, July 2007.

[156] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM. 2000.

[157] F. R. K. Chung, *Spectral Graph Theory.* Providence, RI: American Mathematical Society, 1997.

[158] A. J. Laub, *Matrix Analysis for Scientists and Engineers*. SIAM: Society for Industrial and Applied Mathematics, 2004.

[159] M. Tinati and T. Rezaii, "Multi-target tracking in wireless sensor networks using distributed joint probabilistic data association and average consensus filter," in *Proceedings of the 2009 International Conference on Advanced Computer Control*, Singapore, pp. 51-56, Jan. 2009.

[160] O. Boric-Lubecke, J. Lin, B. Park, C. Li, W. Massagram, V. Lubecke, and A. Host-Madsen, "Battlefield triage life signs detection techniques," in *Proceedings of the SPIE Defense and Security Symposium*, Orlando, FL, vol. 6947, pp. 69470J.1-69470J.10, Apr. 2008.

[161] C. Soto, B. Song, and A. Chowdhury, "Distributed multi-target tracking in a self-configuring camera network," in *Proceedings of IEEE Computer Vision and Pattern Recognition*, Miami, FL, pp. 1486-1493, Jun. 2009.

[162] S. Barbarossa, G. Scutari and A. Swami, "Achieving consensus in self-organizing wireless sensor networks: The impact of network topology on energy consumption," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 2007*, Honolulu, HI, USA, vol. 2, pp. II-841, April 2007.

[163] S. Sardellitti, S. Barbarossa and A. Swami, "Average consensus with minimum energy consumption: optimal topology and power consumption," in *Proceedings of the European Signal Processing Conference*, Aalborg, Denmark, pp. 189-193, Aug. 2010.

[164] M. Nevel'son, and R. Has'minskii, *Stochastic Approximation and Recursive Estimation.* Providence, Rhode Island: American Mathematical Society, 1973.

[165] H. Kushner, and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications.* 2nd ed. New York: Springer, 2003.

[166] N. Ghasemi, S. Dey, and J. Baras, "Stochastic average consensus filter for distributed HMM filtering: Almost sure convergence," Institute for Systems Research Technical Reports, May 2010.

[167] B. S. Y. Rao, H. Durrant-Whyte, and J. A. Sheen, "A fully decentralized multi-sensor system for tracking and surveillance," *The International Journal of Robotics Research*, vol. 12, no. 1, pp. 20-44, Feb. 1993.

[168] B. Mohar, Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, "The Laplacian Spectrum of Graphs," *Graph Theory, Combinatorics, and Applications*, vol. 2, pp. 871-898, 1991.