Control of single and multiple agent systems with input and communication delays

Liu, Shuai

2011

Liu, S. (2011). Control of single and multiple agent systems with input and communication delays. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/48063>

<https://doi.org/10.32657/10356/48063>

Control of Single and Multiple Agent Systems with Input and Communication Delays

LIU SHUAI

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirements for the degree of
Doctor of Philosophy

2011

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

.....

Date

.....

LIU SHUAI

Acknowledgements

I owe my deepest gratitude to my supervisor, Professor Lihua Xie, whose encouragement, professional guidance and support from the initial to the final level enabled me to develop an understanding of the subject. His patient and insightful discussions are greatly appreciated. What I appreciated most were the many opportunities and the freedom he has offered that allowed me to find the research topics I am most interested in. Without his great enthusiastic help, I would not be able to complete this thesis.

I would also like to thank Professor Huanshui Zhang and Professor Frank L. Lewis for their helpful discussions and advice.

I would like to show my gratitude to the members of Sensor Network Lab and other friends who helped and supported me in these years.

I must also acknowledge the research scholarship I received from the Nanyang Technological University when the research work was carried out in the school of Electrical and Electronic Engineering.

In the end, I would like to give my special thanks to my wife Chang Li and my parents for the patient love and encouragement.

Abstract

Control of time-delay systems represents one important class of control problems. A lot of work has been done in recent years. Much of the research work has been focused on the stability analysis and stabilization of time-delay systems based on the Lyapunov functional and linear matrix inequality (LMI) approaches. While the LMI approach does provide an efficient tool, the results are mostly only sufficient and only numerical solutions are available. Our research aims to present analytical solutions by some algebraic approaches, which can provide an insightful understanding of control problem of time-delay systems.

On the other hand, there has also been significant recent interest in networked multi-agent system (MAS) where unavoidably there exist delays in information acquisition as well as information exchange between agents in addition to possible input/state delays in each agent. This motivates the study of MASs with input and communication delays.

The thesis is divided into two parts: Part I is focused on optimal control of single agent systems with multiple input delays; Part II is on distributed consensus control of MAS with communication and input delays.

More specifically, in Part I, we shall study the linear quadratic regulation (LQR) and tracking for a single agent with linear dynamics. We first focus on LQR for linear discrete-time systems with multiple delays in a single input channel. By introducing a backward stochastic system, we establish a relationship between the LQR problem for the original system and optimal estimation for a backward stochastic system

which can be solved by using projection approach. The explicit optimal controllers are developed for both finite time and infinite time horizon cases. A parallel result for linear continuous-time systems is also established, where the optimal feedback gain is given in terms of the solution of Riccati-type partial differential equations (PDEs).

Next, we study the optimal tracking problem for both discrete-time and continuous-time systems with multiple delays in a single input channel. Tracking problem requires the states and the input to track desired trajectories rather than converge to the origin. We convert the linear quadratic tracking problem into a white noise and state smoothing problem for an associated stochastic backward system. A deterministic backward delay system is introduced for establishing an explicit expression of the optimal controller.

In Part II, we focus on distributed control of MASs with communication and input delays. We first comparatively study consensus protocols with and without control inputs information of neighbors. We will rigorously prove that knowing neighbors' control inputs information can improve the consensus performance. However, the protocol with neighbors' control inputs information cannot be implemented in practical systems. Therefore, a distributed protocol based on neighbors' delayed control input information is proposed and investigated. We also investigate how the time delay affects the consensus performance.

Next, we study distributed consensus for MAS with both delays and noises in transmission channels. Due to the random transmission noises, traditional protocol cannot be applied. By introducing a decaying consensus gain, we can provide necessary and sufficient conditions on the consensus gain and communication graph to guarantee that the whole system reaches the consensus in a stochastic sense.

In a digital network, quantization is unavoidable. However, quantization error cannot be treated as a random transmission noise in most of cases. Therefore, we study distributed consensus for MAS with communication delays and limited data rate. A

distributed consensus protocol is proposed based on dynamic encoding and decoding. It will be shown that for a connected network, as long as the time delays are bounded, the average consensus can be achieved with a finite communication data rate. Moreover, merely one bit information exchange between each pair of adjacent agents at each time step suffices to guarantee the average consensus.

Symbols and Acronyms

Algebraic Operators

$[a, b]$	closed interval from a to b
A'	transpose of A
A^{-1}	inverse of matrix A
$col\{x_1, \dots, x_n\}$	column vector formed by $x_1, x_2, \dots, x_n$
$diag\{x_1, \dots, x_n\}$	block diagonal matrix formed by $x_1, x_2, \dots, x_n$
$\langle x, y \rangle$	mathematical expectation of xy'
$\langle x, y \rangle_{\mathcal{H}}$	inner product over Hilbert space $\mathcal{H}$
$\delta_{i,j}$	Kronecker delta function
$\delta(t)$	Dirac delta function
$\chi_{\Omega}(t)$	indicator function, i.e. $\chi_{\Omega}(t) = \begin{cases} 1, & \text{if } t \in \Omega \\ 0, & \text{if } t \notin \Omega \end{cases}$
$\mathcal{L}\{\cdot\}$	linear space spanned by the elements in the braces
$\hat{\mathbf{x}}(t_1 t_2)$	projection of $\mathbf{x}(t_1)$ onto the linear space spanned by measurements up to time instant t_2
$\tilde{\mathbf{x}}$	projection error, i.e. $\mathbf{x} - \hat{\mathbf{x}}$
$\sigma(A)$	spectrum set of A
$\rho(A)$	resolvent set of A
$O(\cdot)$	equal-order or higher order infinitesimal, i.e. $f(\tau) = O(\tau)$ if and only if $\lim_{\tau \rightarrow 0} \left \frac{f(\tau)}{\tau} \right < +\infty$
$A^{i,j}$	the element in the i -th row and j -th column of matrix A
I_n	$n \times n$ identity matrix

$\mathbf{1}_n$	an n -dimensional column vector with all elements of 1
$\mathbf{0}_n$	an n -dimensional column vector with all elements of 0
$\text{Re}(\cdot)$	the real component of a complex number
$\ \cdot\ _2$	Euclidean norm
$\ \cdot\ _\infty$	infinity norm
$m_1 \wedge m_2$	the smaller number of m_1 and m_2 , i.e. $\min\{m_1, m_2\}$
$m_1 \vee m_2$	the larger number of m_1 and m_2 , i.e. $\max\{m_1, m_2\}$
$e_{m,i}$	an m -dimensional column vector with the i -th entry being 1 and all the other entries 0
$E(X)$	the mathematical expectation of X
$\text{Var}(X)$	the variance of X
$P(\mathcal{A})$	the probability of event $\mathcal{A}$
$\Pi_{i,j}^M$	transition matrix generated by M , i.e. $\Pi_{i,j}^M = \begin{cases} M(i)M(i-1) \cdots M(j), & i \geq j \\ I, & i < j \end{cases}$
$\lceil \cdot \rceil$	ceiling function
$\lfloor \cdot \rfloor$	floor function

Sets and Spaces

$\mathbb{C}$	set of complex numbers
$\mathbb{R}$	set of real numbers
$\mathbb{R}^+$	set of positive real numbers
$\bar{\mathbb{R}}^+$	set of nonnegative real numbers
$\mathbb{R}^n$	Set of n dimensional real vectors
$\mathbb{R}^{m \times n}$	set of $m \times n$ real matrices
$\mathbb{Z}$	set of integers
$\mathbb{Z}^+$	set of positive integers
$\bar{\mathbb{Z}}^+$	set of nonnegative integers

$L_2(a, b; \mathbb{R}^n)$	space of all equivalent classes of square integrable n -dimensional real vector functions on $[a, b]$
$W^1(a, b; \mathcal{X})$	Sobolev space of all functions x in $L_2(a, b; \mathcal{X})$ with first order derivative

Acronyms

a.s.	almost surely
ARE	algebraic Riccati equation
LMI	linear matrix inequality
LQR	linear quadratic regulation
LQG	linear quadratic Gaussian
MPC	model predictive control
MAS	multi-agent system
PDE	partial differential equation
UGV	unmanned ground vehicle
UAV	unmanned aerial vehicle

Contents

Acknowledgement	i
Abstract	iii
Symbols and Acronyms	vii
List of Figures	xv
1 Introduction	1
1.1 Motivation and Objectives	1
1.2 Major Contributions	4
1.3 Outline of the Thesis	5
2 Literature Review	9
2.1 Optimal Control of Single Agent Time-delay Systems	9
2.2 Consensus Control of MASs with Time Delays	13

Part I Optimal Control of Single Agent Systems with Input Delays **19**

3 LQR for Discrete-time Systems with Multiple Delays in a Single Input Channel	21
3.1 Problem Statement	22
3.2 Solution to the Discrete-time LQR Problem	23
3.2.1 Duality between LQR and Smoothing	24
3.2.2 Solution to the Finite Horizon LQR Problem	28
3.2.3 Solution to the Infinite Horizon LQR Problem	33
3.3 Numerical Examples	37
3.4 Conclusion	38
 4 LQR for Continuous-time Systems with Multiple Delays in a Single Input Channel	 41
4.1 Problem Statement	42
4.2 Solution to the Continuous-time LQR Problem	43
4.2.1 Duality between LQR and Smoothing	44
4.2.2 Solution to the Finite Horizon LQR Problem	47
4.2.3 Solution to the Infinite Horizon LQR Problem	53
4.3 Conclusion	59

5	Linear Quadratic Tracking for Discrete-time Systems with Multiple Delays in a Single Input Channel	61
5.1	Problem Statement	62
5.2	Main Result	63
5.3	Proof of Main Result	65
5.3.1	Duality between Tracking and Smoothing Problems	66
5.3.2	Formation of the Backward System and Solution to the Tracking Problem	71
5.4	Analysis of Computational Complexity	75
5.5	Applications	77
5.6	Conclusion	78
6	Linear Quadratic Tracking for Continuous-time Systems with Multiple Delays in a Single Input Channel	81
6.1	Introduction	81
6.2	Problem Statement	82
6.3	Main Result	83
6.4	Proof of Main Result	85
6.4.1	Duality between Tracking and Smoothing Problem	85
6.4.2	Formation of the Backward System and Solution to the Tracking Problem	89
6.5	Conclusion	92

Part II Distributed Control of Multi-agent Systems with Communication/Input Delays 95

7 Distributed Consensus for MAS with Delayed Control Input Information from Neighbors	97
7.1 Problem Statement	98
7.2 Main Results	101
7.2.1 Comparison	101
7.2.2 Protocol with Delayed Control Input from Neighbors	107
7.2.3 Application to Formation Control	119
7.3 Simulations and Experiments	121
7.3.1 Consensus Problem	121
7.3.2 Formation Problem	125
7.4 Conclusion	130
8 Distributed Consensus for MAS with Delays and Noises in Transmission Channels	133
8.1 Problem Statement	134
8.2 Distributed Consensus	136
8.2.1 Fixed Topology	136
8.2.2 Dynamically Switching Topologies	157
8.2.3 Randomly Switching Topologies	160
8.3 Numerical Examples	164
8.4 Conclusion	167
8.5 Appendix: Proof of Lemma 8.4	169

9 Distributed Consensus for MAS with Communication Delays and Limited Data Rate	177
9.1 Problem Statement	178
9.2 Protocol Design and Consensus Analysis	179
9.2.1 Protocol Design	179
9.2.2 Consensus Analysis	181
9.3 Numerical Examples	201
9.4 Conclusion	202
9.5 Appendix: Proof of Lemma 9.2	203
10 Conclusions and Future Work	207
10.1 Conclusions	207
10.2 Future Work	209
Appendix	211
A Preliminaries in Graph Theory	211
Author's Publications	213
Bibliography	215

List of Figures

3.1	State feedback control	38
3.2	Trajectory of state	39
5.1	Comparison of complexities between augmentation method and the presented approach	76
5.2	Sales of products.	79
5.3	Inventory of products.	79
5.4	Production of products.	80
5.5	Advertisement of products.	80
6.1	Water distribution system	82
7.1	Augmented Graph $\bar{\mathcal{G}}$	122
7.2	Dot line: state trajectories under protocol (7.6); solid line: state trajectories under protocol (7.5)	122
7.3	Dot line: $\tau = 0s$; solid line: $\tau = 0.2s$	123
7.4	Dot line: $\tau = 0s$; solid line: $\tau = 2s$	124
7.5	Dot line: $\tau = 0s$; solid line: $\tau = 0.2s$	124
7.6	Dot line: $\tau = 0s$; solid line: $\tau = 2s$	125

7.7	Dot line: Trajectories of the two agents; solid line: Trajectory of reference state	125
7.8	x coordinates of the two agents and reference state	127
7.9	y coordinates of the two agents and reference state	127
7.10	Trajectories of the two agents and the reference state	128
7.11	Model of Amigo robot	128
7.12	x coordinates of the Amigos given static reference state (7.54)	129
7.13	y coordinates of the Amigos given static reference state (7.54)	130
7.14	Formation by Amigo given dynamic reference state (7.55)	130
8.1	Quotient graph of the MAS.	165
8.2	Leader-follower topology with $c(k) = 0.5$	166
8.3	Leader-follower topology with $c(k) = 1/\sqrt{k + \bar{d}}$	166
8.4	Fixed undirected topology with $c(k) = 1/\sqrt{k + \bar{d}}$	167
8.5	Fixed undirected topology with $c(k) = 1/(k + \bar{d})$	168
8.6	Dynamically switching topologies with $c(k) = 1/(k + \bar{d})$	168
8.7	Randomly switching topologies with $c(k) = 1/(k + \bar{d})$	169
9.1	The transmission channel from i to j	179
9.2	The encoder ϕ_i	180
9.3	The decoder $\varphi_{i,j}$	181
9.4	State trajectories of 5 agents with $h = 0.0005$ and $\gamma = 0.9993$	201
9.5	State trajectories of 5 agents with $h = 0.01$ and $\gamma = 0.99$	202

Chapter 1

Introduction

1.1 Motivation and Objectives

In many control systems, for example, transportation systems, communication networks, chemical processes, bio-systems, economics, process engineering and more recently networked control systems, there exist situations where the process depends not only on the present state and control input, but also on their past information. They correspond to measurement time, transmission time and computation times. In such cases, the previous information is involved in the system evolution. These systems are called time-delay systems. Time-delay systems widely exist in many engineering fields. One such example is steel rolling process where the measurement point is located at some place downstream of the steel press so that there is always a time delay in the feedback of plate thickness to the press controller [124].

Numerous investigations have shown that temporal delays in a system have a considerable influence on the qualitative behavior of the system. Many phenomena, such as periodicity, oscillation, instability, etc., can be explained in terms of delay. The field of time-delay systems had its origin in the 18th century, and it received substantial attention in the early 20th century in works devoted to the modeling of biological, ecological, and engineering systems [39]. Time delay cannot be ignored

otherwise the closed loop system performance may not be guaranteed, even worse, stability may be jeopardized. Sometimes an arbitrarily small delay in feedback regulator may radically change the qualitative behavior of the system. For example, consider the dimensionless Euler-Bernoulli equation of transversal oscillation of a beam^[60]:

$$\frac{\partial^2 \omega(t, x)}{\partial t^2} + \frac{\partial^4 \omega(t, x)}{\partial x^4} = 0, \quad 0 \leq x \leq 1, \quad t > 0,$$

with boundary conditions

$$\omega(t, 0) = \frac{\partial \omega(t, 0)}{\partial x} = \frac{\partial^3 \omega(t, 1)}{\partial x^3} = 0$$

and feedback

$$\frac{\partial^2 \omega(t, 1)}{\partial x^2} = -\frac{\partial^2 \omega(t - h, 1)}{\partial t \partial x}.$$

Here, $\omega(t, x)$ is the displacement of the beam at point x and time t . It is known that for $h = 0$ the Euler-Bernoulli beam is uniformly exponentially stabilized by the given feedback. However (see [22]), the system becomes unstable if an arbitrarily small delay $h > 0$ is introduced into the feedback, and arbitrarily large divergence rates are generated.

Note that designing a controller for a time-delay system is usually complicated. Furthermore, in industry applications, we often require not only the stability of system, but also certain optimal performance, such as minimal energy consumption. The study of optimal control for systems with delays has a history of over forty years. The linear quadratic regulation was a hot topic in the early development. However, it should be noted that the linear quadratic regulator is hard to be designed for many time-delay systems especially for systems with multiple delays in a same transmission channel. It is because that the signal not only affects the current system but also the system in the future. The coupled effects make the controller design much more complicated.

On the other hand, with the emergence of networked control systems, multi-agent systems replace a single agent system to execute more challenging tasks or some

tasks which cannot be accomplished by a single agent. Therefore, the research on cooperative control of MASs has received more and more attention. There is no universal definition for cooperative control. It can be understood as a collection of decision-making agents with limited processing capabilities, locally sensed information, and limited inter-agent communications, all seeking to achieve a collective objective [104]. As an application, cooperative control is used to solve the problems of consensus, rendezvous, tracking of a virtual leader, formation control and synchronization [95].

Distributed consensus is an important problem closely related to cooperative control. In consensus problem, information is collected through sensing networks, where transmission delays often exist. When we design a consensus algorithm, the existence of delays in transmission channel, measurement process and computation should be taken into consideration. There are some works on consensus problem with communication/input delays. It has been proved that consensus can be achieved in the presence of bounded communication delays. As for input delays, there is an upper bound to guarantee consensus. However, for MASs with communication/input delays, there are many consensus problems left open. For example, when we consider measurement noises or transmission noises, consensus problems become more complicated and traditional consensus protocol does not work for communication topologies with delays.

The research in this thesis includes optimal control of single agent systems with multiple input delays and distributed consensus control of multi-agent systems with communication/input delays. The objectives to be achieved are: 1) Present analytical solutions to optimal control problems for systems with multiple delays in a single input channel. Based on that, more complicated and more general control problems for input delay systems will be investigated; 2) Design distributed consensus protocols for MASs with communication and input delays. Some complex systems, e.g. MASs with transmission noises or communication data rate constraint will be studied.

1.2 Major Contributions

The main contributions of the thesis are listed as follows:

1. We design a linear quadratic regulator for linear systems with multiple delays in a single input channel. The results are extended to infinite time horizon case.
2. We solve the linear quadratic tracking problem for linear systems with multiple delays in a single input channel. The optimal controller is given in terms of the current state, the previous control inputs and the output of an auxiliary backward deterministic delay system.
3. We comparatively study two control protocols – the protocols with and without neighbors' control input information. We show that the latter can lead to a better consensus performance. We propose a new consensus protocol for MASs by introducing a time delay in the neighbors' control input information to make the protocol implementable. The effects of the introduced delay on the consensusability and consensus speed are analyzed.
4. We study the consensus problem for linear discrete-time MASs with delays and noises in transmission channels. A novel technique is introduced to overcome the difficulties induced by the delays and noises. Necessary and sufficient conditions for strong consensus and mean square consensus are respectively given for non-leader-follower and leader-follower cases under a fixed communication topology. Under dynamically switching topologies and randomly switching topologies, sufficient conditions for strong consensus and mean square consensus are also obtained.
5. We consider the average consensus problem for multi-agent networks with communication delays and limited data rate. A distributed consensus protocol is proposed based on dynamic encoding and decoding. We show that the average consensus can be achieved with finite communication data rate

under the proposed protocol. Moreover, merely one bit information exchange between each pair of adjacent agents at each time step suffices to guarantee the average consensus.

1.3 Outline of the Thesis

Some related works will be reviewed in the following chapter. The remaining chapters are organized into two parts. The first part (including Chapter 3-Chapter 6) deals with control of single agent systems with multiple input delays while the second (including Chapter 7-Chapter 9) focuses on distributed consensus control of multi-agent systems with communication and input delays.

In Chapter 3, we will consider linear quadratic regulation for discrete-time systems with multiple delays in a single input channel. By introducing an auxiliary backward stochastic system and through re-organization of output measurements, the deterministic quadratic optimization problem is converted into a smoothing estimation problem. A new controller derivation method is then developed which overcomes the difficulty posed by multiple delays in a single input channel. Our solution to the LQR problem is given in terms of the solution of a two-dimensional Riccati difference equation.

In Chapter 4, a parallel result of linear quadratic regulation for continuous-time systems with multiple delays in a single input channel will be studied. Different from discrete-time systems, it is well known that linear continuous-time systems with time delays are infinite dimensional systems. State augmentation approach does not work. Based on re-organization approach, the optimal control problem is converted into a smoothing problem for a backward stochastic system. Analytical results are obtained.

In Chapter 5, we will extend the LQR result of Chapter 3 to linear quadratic tracking for discrete-time systems with multiple delays in a single input channel. Tracking problem is more difficult than linear quadratic regulation since it requires the states

to track a desired trajectory rather than converge to the origin. By applying the re-organization approach, the optimal tracking problem is successfully solved.

In Chapter 6, we will study linear quadratic tracking for continuous-time systems with multiple delays in a single input channel. The result for linear quadratic regulation is extended to the linear quadratic tracking problem. The optimal feedback controller is given in terms of the solution of partial differential equations and a backward deterministic system.

In Chapter 7, we will design a distributed consensus protocol for multi-agent systems with delayed control input information from neighbors. We comparatively study two control protocols, a standard local voting protocol, and a protocol that includes information about the control inputs of neighbors. It will be shown that using control input information from neighbors improves the consensus performance. Realistically, control input information from neighbors can only be received after a time delay. It will be therefore shown that the second protocol using time-delayed control input information from neighbors guarantees consensus of the multi-agent system for any nonnegative delay. Furthermore, the convergence rate using time-delayed control inputs can still be selected with considerable freedom, and remains superior to the performance of the standard local voting protocol which depends on the graph topology.

In Chapter 8, we will investigate distributed consensus for multi-agent systems with delays and noises in transmission channels. Due to the presence of noises and delays, existing techniques such as the lifting technique and the stochastic Lyapunov theory are no longer applicable to the analysis of consensus. A novel technique is introduced to overcome the difficulties induced by the delays and noises. A consensus protocol with decaying gains satisfying persistence condition is adopted. Necessary and sufficient conditions for strong consensus and mean square consensus are respectively given for non-leader-follower and leader-follower cases under a fixed topology.

In Chapter 9, we will investigate distributed consensus for multi-agent systems with communication delays and limited data rate. The average consensus problem becomes much more complicated when both delays and data-rate constraint are to be

considered. A distributed consensus protocol is proposed based on dynamic encoding and decoding. It is shown that for a connected network, as long as the time delays are bounded, the average consensus can be achieved with a finite communication data rate. In particular, merely one bit information exchange between each pair of adjacent agents at each time step suffices to guarantee the average consensus.

Finally, Chapter 10 draws some concluding remarks. The contributions of thesis are recaptured and possible future research topics are highlighted.

Chapter 2

Literature Review

Control problems for time-delay systems were first studied in [11] and [41]. Lyapunov's second method for the stability of delay systems was developed at the end of the 1950s by N. Krasovskii who introduced Lyapunov functionals and by B. Razumikhin who proposed appropriate Lyapunov functions [34]. Since the 1950s, the subject of delay systems or functional differential equations has received a great deal of attention in Mathematics, Biology and Control Engineering. The first effective control scheme for time-delay systems was originally introduced in [106] and [107]. By introducing a minor feedback loop consisting of a delay predictor, the controller design problem for a time-delay system is then converted to a design problem for a delay-free system. The output responses after the duration of delay can be adjusted arbitrarily. This approach motivated the finite-spectrum assignment [76]. Considerable research has been done on various aspects of time-delay systems from the 1960s.

2.1 Optimal Control of Single Agent Time-delay Systems

Since the late 1960s, optimal control of time-delay systems has been an active research area and a lot of results have been put forward to solve the optimal control

problem of time-delay systems in some sense, see e.g. [21, 35, 113, 134] for the discrete-time case and [29, 1, 23, 121, 112, 79] for the continuous-time case. In the early days, the LQR was the main research topic. It should be noted that the closed-loop system with a linear quadratic optimal controller has desirable sensitivity and robustness properties.

For continuous time-delay systems, the control problem is often treated by the infinite-dimensional system theory approach [8] and the linear quadratic optimal control design often requires solutions to infinite dimensional Riccati equations or operator Riccati equations. Some of the Riccati equations can be converted into finite dimensional Riccati differential equations. The solution for the LQR problem is provided in [122, 23, 94]. In [122, 49], a family of evolution equations is introduced which transforms the input delay systems into delay-free ones. Under the infinite dimensional systems setting, Delfour [23] studies continuous-time systems with a general delay structure in both state and input and presents a solution based on a Riccati partial differential equation. In [79], an LQR problem with multiple input delays is considered. The optimization time interval is split into time regions whose boundaries are the delays and dynamic programming can be used within each time region.

The H_∞ control problem has been well investigated in [114, 56, 112, 58]. Applying linear transformation, the stabilization problem for systems with single input delay is investigated in [33]. Elegant solutions to the H_∞ control of systems with single input delay have been presented in [112] and [140]. By converting the delay problem into a nested sequence of elementary problems, Meinsma and Mirkin [77] give a complete solution to the H_∞ control problem for systems with multiple I/O delays. The solvability and explicit solution depend on a number of J -spectral factorizations. Further, Kojima and Ishijima [58] investigate the H_∞ preview control of systems with multiple input delays. Since the general framework of infinite-dimensional systems is used, the solutions are very complicated and difficult for engineers to understand.

Based on the duality principle, Krein space approach has long been recognized and is well presented in [42]. Given any quadratic cost function, a Krein space model related to this function is constructed, for which the solution to the optimal control problem is given in terms of the negative transpose (adjoint) of the operator that yields the Krein space projection. By introducing the reorganization approach [137, 138], the Krein space approach is generalized and extended to time-delay systems. To be specific, Zhang and Xie [135] convert the LQR problem for a continuous-time system with multiple input delays into an estimation problem for an associated stochastic backward system (Krein space model) with multiple output delays. The reorganization approach presents simple analytical solutions to control and estimation problems for systems with multiple I/O delays.

On the other hand, there has been increasing interest in discrete-time delay systems in recent years due to applications in the fields of networked control and network congestion control. Augmentation is a straightforward method to convert the delay system into delay-free one, but it leads to high state dimension and thus high computational cost. Based on a duality principle, a new way to deal with the control problem for systems with input delay is provided in [134] and [135]. The duality allows one to address complicated multiple input delay problems via some elementary tools such as projection. However, we note that in [134, 135] only finite horizon control problems have been investigated.

More and more practical systems are modeled as systems with multiple delays in a single input channel recently. Lian *et al.* [67] discuss the modeling and analysis of networked control systems where sensors, actuators, and controllers are distributed and interconnected by a common communication medium. Therefore, multiple distributed communication delays as well as multiple inputs and multiple outputs in a single channel appear in the model. However, there are very few works on control problem for such systems due to its inherent difficulty. The challenge in control problems for systems with multiple delays in a single input channel is the interaction between the input signals at different time instants. The effect of control input on

both current and future states should be taken into account simultaneously when the controller is designed.

It is noted that control problems for systems with single input but multiple delays which contain the systems studied in [135] as a special case becomes a constrained optimization problem if the approach of [135] is applied directly. Though state augmentation method is straightforward for discrete-time systems, it will lead to high state dimension and thus high computational cost. Furthermore, the augmentation method is unavailable for continuous-time system. Ichikawa [49] provides a general theory for a family of evolution equations with a control operator containing a finite number of pure delays. Delfour [23] generalizes the results of [122] to a control operator with a more general delay structure than the one in [49]. These results are based on infinite dimensional system theory, which is very difficult to understand.

Different from regulation, tracking problem is concerned with the design of a control law such that the output of a system under investigation is able to follow a given reference input. The tracking problem for time-delay systems has been studied in the past few decades e.g., [92], [54], [71], [123], [117].

For discrete-time delay systems, the tracking problem can be solved by applying a state augmentation method, see [20], [54]. In [54], the dead beat control approach is applied to minimize the convergence time. However, due to the above-mentioned disadvantages of the augmentation approach, approaches without resorting to state augmentation are much preferred. In [92], a linear quadratic tracking problem for systems with single input delay is solved based on the minimum principle and the result is applied to economic systems in [91]. A similar problem has been studied by [123] where an iterative algorithm is proposed to calculate an approximate optimal controller by the Artstein model reduction method. For discrete-time delay systems, by using successive approximation approach, Tang *et al.* [118] investigate the suboptimal tracking problem for systems with state delay. For systems with multiple input channels and signal delay in each channel, Liu *et al.* [71] presents a way to calculate the optimal tracking controller based on a duality principle.

2.2 Consensus Control of MASs with Time Delays

Recently, distributed cooperative control has been attracting a lot of attention by many researchers. Distributed cooperative control, compared with centralized control, has many advantages, such as energy saving, delay robustness, fault tolerance, scalability, etc. Consensus problems for MASs, whereby all the agents will reach a prescribed reference state asymptotically or in a finite time, are important problems when studying distributed cooperative control, see [98, 104, 96], etc. For some consensus problems, the common value to which the states converge is also required to be the average of the initial states of the group. It is called average-consensus.

Many seemingly different problems that involve interconnection of dynamic systems in various areas of science and engineering happen to be closely related to consensus problems for MASs [85].

- **Flocking and rendezvous**

A common form of consensus problem is rendezvous in space [68], [27], [109]. This is equivalent to reaching a consensus in position by a number of agents with an interaction topology that is position induced. Flocking is somewhat more challenging than rendezvous in space because it requires both inter-agent and agent-to-obstacle collision avoidance. Flocks of mobile agents equipped with sensing and communication devices can serve as mobile sensor networks for massive distributed sensing in an environment. The role of consensus algorithms in particle based flocking is for an agent to achieve velocity matching with respect to its neighbors [51].

- **Formation control**

Formation control for collections of vehicles has many advantages [32], [63], [108]. Using the formation control results in better stability, robustness and energy-saving. For example, in small satellite orbiting, formation helps to

reduce the energy consumption, expand their sensing capabilities and decrease targets access time. In an automated highway system, which is an advanced intelligent transportation system technology designed to provide for driverless cars, the throughput of the transportation network can be greatly increased if vehicles can form to platoon at a desired velocity while keep a specified distance between vehicles.

- **Synchronization of coupled oscillators**

The problem of synchronization of coupled oscillators has attracted numerous scientists from diverse fields including physics, biology, neuroscience, and mathematics [98], [78]. This is partly due to the emergence of synchronous oscillations in coupled neural oscillators.

- **Distributed sensor fusion**

The most recent application of consensus problems is distributed sensor fusion in sensor networks. This is done by posing various distributed averaging problems required to implement a Kalman filter [87], [57], [84], low-pass and high-pass consensus filters [111], linear least-squares estimator [130] as average-consensus problems.

For the Vicsek's model where all the agents move in a plane at the same speed but with different headings, Jadbabaie *et al.* [51] provide a simple consensus protocol which updates the heading of an agent by averaging the headings of its neighbors. The protocol guarantees that all the agents can move in the same direction eventually. Ren [100] develop a protocol with bounded control inputs by introducing $\tanh(\cdot)$ function. Using algebraic theory and graph theory, Cao *et al.* [12] provide a systematic way to analyze a class of consensus problems under dynamically changing environments. It gives a lower bound of consensus convergence rate. Readers who are interested in the convergence rate problem are referred to [120], [4], [88], [141].

Most of the existing literature on consensus is about linear systems. Recently, nonlinear consensus problems have also been investigated. To reach a consensus on a nonlinear function of the initial values, Bauso *et al.* [7] introduce a non-linear protocol design rule. In the case when the plant is nonlinear, the results on the consensus problem for a wide class of continuous-time and discrete-time systems are provided in [3] and [70] (generalized results from [80]), respectively. Under the assumption that the vector fields satisfy a certain tangentiality condition, a necessary and sufficient condition for consensus is that the graphs are jointly connected.

In the stochastic setting, the consensus problem has been studied in the presence of random link failures and/or transmission noises. For the case of random link failures, Fagnani and Zampieri [31] provide two kinds of average consensus protocols which are biased compensation method and balanced compensation method. The performance degradation due to the link failures is evaluated. The corresponding consensus convergence rate is investigated in [141]. Under the traditional consensus protocol, Pereira and Pages-Zamora [90] analyze the mean square consensus error and studies the optimal consensus gain problem.

On the other hand, noises often come from measurement errors, environment disturbances during transmission and quantization errors. In [132], it is shown that in the presence of noises the consensus value will diverge when the traditional consensus algorithms are applied. By removing the common assumption that all the nonzero weights are uniformly bounded from below, Huang and Manton [45] develop a stochastic approximation type of algorithm to deal with noises. Under a fixed topology, Li and Zhang [66] obtain necessary and sufficient conditions for mean square average consensus. A sufficient condition is also given for the switching topologies case [65], [44]. The asymptotic normality property of the difference between the state vector and its limiting value for each agent is established in [47]. Kar and Moura [53] study the average consensus with both randomly broken links and channel noises. Two different algorithms are provided. The first one is to force the weights to satisfy the so-called persistence condition, which is the same as that of [45], [66]. The second one is to run the consensus iterations for $\hat{\ell}$ steps and repeat

the procedure $\hat{p}$ times. The average of the $\hat{p}$ results of the iterations reaches an (ϵ, δ) consensus.

In digital communication networks, communication channels only have finite capacities, that is, an agent can only transmit a finite bits of information to its neighbors at each time step. Therefore, consensus with quantized communication has drawn the attention of the systems and control community. A gossip-algorithm-based approximate average consensus protocol is provided by [55] under the assumption that the states of agents are integer-valued. Under the same assumption, Nedic *et al.* [82] analyze the quantization effect on average consensus and give an upper bound for the consensus error. In [17], three kinds of update strategies are considered based on both deterministic and probabilistic quantizers. The average consensus problem is solved by using logarithmic quantizers with infinite level, see [15] and [16]. Under the assumption that the quantization errors are white noises, two coding schemes are provided in [133] and conditions under which the consensus can be achieved are obtained. By introducing dynamic encoding and decoding schemes, it is proved in [64] that the average consensus problem can be solved by using only finite level quantizers. Moreover, Li *et al.* [64] prove that the number of transmitting bits at each time step can be reduced to one by properly selecting the controller parameters.

In practical networked MASs, delays are unavoidable in information acquisition and transmission which should be taken into consideration when designing a consensus protocol. For MASs with delays, the consensus problems become complicated [86, 128, 119, 13, 10, 74, 81]. The delays can be categorized into two classes. In class I, for each agent delays only exist in the acquisition of neighbors' information. In class II, delays exist in both neighbors' and its own information. It contains a variety of delays, such as input delays, communication delays and measurement delays.

For class I, a synchronization criterion for coupled oscillators with uniform communication delays is given in [28]. Under the traditional consensus protocol, Cao *et al.* [13] analyze the consensusability and consensus speed for MASs with first order discrete-time dynamics. The consensus rate is investigated by [73]. For MASs

with second order dynamics, consensus problem is investigated by [69], in which time delays happen in the acquisition of both the position and velocity of a neighbor. Based on a Lyapunov approach, the robustness of the traditional consensus protocol against time delay for MASs with first order continuous-time dynamics is studied by [74]. The result is also extended to MASs with nonlinear dynamics. The effects of delay in multi-agent oscillator synchronization is studied in [89]. It is concluded that the consensus can still be guaranteed by applying traditional consensus protocol for arbitrary delays in class I.

For class II, traditional consensus protocol however may fail to work, see [119, 128, 10, 81]. Olfati-Saber and Murray [86] give an initial study on the consensus problem for continuous-time systems in frequency domain and provides a necessary and sufficient condition on the upper bound of time delays under the assumption that all the delays are equal and time-invariant. Bliman and Ferrari-Trecate [10] recast the continuous-time systems with transmission delays and provides a systematical analysis on average consensus problem. A necessary and sufficient condition for the existence of average consensus under invariant communication delays (same as that in [86]) is derived. The upper bound of time delays for nonuniform delays and even time-varying delays to guarantee the consensus for first order continuous-time MASs is also analyzed. For discrete-time MASs, Tian and Liu [119] provide a sufficient condition on the consensus gain. Similar problem is studied by [128], in which a new consensus protocol is introduced which is more general than the existing ones. A generalized Nyquist consensus condition is presented in [81] for MASs of a general dynamic model and different feedback delays. For conventional protocols, it is shown that delay may slow down the consensus process [13], [83]. Nevertheless, delayed information can have positive effects on convergence rate when they are properly used [59], [14].

For more complicated problems, e.g. when transmission channel suffers from both delays and noises, or even data rate constraint, traditional consensus protocol cannot work. The lifting technique developed in [128] and [13] to deal with delay problems cannot be directly applied when we consider noises in transmission channel [132].

The stochastic Lyapunov theory developed in [46] and [66] to deal with transmission noises is no longer valid because the existence of delays leads to that no such stochastic Lyapunov function can be found. Those complicated consensus problems will be investigated in this thesis.

Part I

Optimal Control of Single Agent Systems with Input Delays

Chapter 3

LQR for Discrete-time Systems with Multiple Delays in a Single Input Channel

In this chapter, we consider the LQR problem for linear discrete-time systems with multiple delays in a single input channel. This kind of systems can be found in process engineering, communication systems with multiple path fading channels, biological systems, economical systems, and networked control systems with delay of more than one sampling period. Using a re-organized output approach, we decouple the input signals at various delayed time instants and establish a duality between the LQR problem and a smoothing problem for an associated backward stochastic system. An analytical solution to the LQR problem is then derived by a new controller derivation method. The solution is given in terms of a Riccati recursion of the same order of the plant (ignoring the delays). Compared with the augmentation approach, our result is computationally much more efficient.

We also solve the infinite horizon LQR problem. Note that the finite horizon results can not be directly extended to the infinite horizon case without establishing the convergence and stability of the closed-loop system. Conditions for convergence and stability are provided in this chapter.

3.1 Problem Statement

We consider the following discrete-time linear system with multiple delays in a single input channel

$$x(k+1) = \Phi(k)x(k) + \sum_{i=0}^l \Gamma_{(i)}(k)u(k-d_i), x(0) = x_0, \quad (3.1)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input with initial values $u(k) = \mu_k$ when $-d_l \leq k < 0$, $\{d_i \in \bar{\mathbb{Z}}^+, i = 0, \dots, l\}$ is the set of constant input time delays, $\Phi(k)$ and $\Gamma_{(i)}(k)$ are bounded time-varying matrices with appropriate dimension. Without loss of generality, the delays are assumed to be of an increasing order: $0 = d_0 < d_1 < \dots < d_l$.

Consider the following quadratic performance index for the system (3.1):

$$J_N = x'_{N+1}P_{N+1}x_{N+1} + \sum_{i=0}^N [u'(i)R_i u(i) + x'(i)Q_i x(i)], \quad (3.2)$$

where $N \geq d_l$ is an integer, x_{N+1} is the terminal state, i.e., $x_{N+1} = x(N+1)$, $P_{N+1} \geq 0$ is the penalty weighting matrix on the terminal state, $R_i > 0, Q_i \geq 0, i = 0, \dots, N$ are weighting matrices on the input signal $u(i)$ and state $x(i)$, respectively.

The finite horizon LQR problem is to find the state feedback control $\{u^*(s), 0 \leq s \leq N\}$ such that the cost function J_N is minimized. We also deal with the infinite horizon case, i.e. the case when $N \rightarrow \infty$.

Remark 3.1. *In this chapter, we first introduce a dual system and convert the LQR problem into a smoothing problem for the dual system. Then an analytical solution to the LQR problem is obtained based on the duality. In our study, we assume that the systems under investigation are without uncertainty and the time delays are time-invariant and known. These allow us to concentrate on the fundamental optimal control problem and provide an analytical solution to the problem. We note that there are existing approaches for systems with uncertain delays where only sufficient*

suboptimal conditions are given and numerical optimization is usually involved to compute control laws.

In the following, we shall solve the LQR problem for the discrete-time systems (3.1) for both the finite and infinite horizon cases. We shall also establish stability and convergence properties of the LQR control.

3.2 Solution to the Discrete-time LQR Problem

We shall convert the discrete-time LQR problem into the optimal smoothing problem for an associated stochastic backward system.

Before introducing the backward stochastic system, we shall introduce some notation. Given $k \geq s \geq 0$, denote

$$u_s(k) \triangleq \begin{cases} \text{col}\{u(k), u(k-d_1), \dots, u(k-d_i)\}, & d_i \leq k-s < d_{i+1}, \\ \text{col}\{u(k), u(k-d_1), \dots, u(k-d_l)\}, & k-s \geq d_l, \end{cases} \quad (3.3)$$

$$\Gamma_s(k) \triangleq \begin{cases} [\Gamma_{(0)}(k) \ \Gamma_{(1)}(k) \ \dots \ \Gamma_{(i)}(k)], & d_i \leq k-s < d_{i+1}, \\ [\Gamma_{(0)}(k) \ \Gamma_{(1)}(k) \ \dots \ \Gamma_{(l)}(k)], & k-s \geq d_l, \end{cases} \quad (3.4)$$

$$\tilde{u}_s(k) \triangleq \sum_{j=0}^l \Gamma_{(j)}(k) u(k-d_j) \chi_{[-d_l, s)}(k-d_j), \quad (3.5)$$

$$R_s(k) \triangleq \begin{cases} \text{diag}\{R_{0,k}, \dots, R_{i,k}\}, & d_i \leq k-s < d_{i+1}, \\ \text{diag}\{R_{0,k}, \dots, R_{l,k}\}, & k-s \geq d_l, \end{cases} \quad (3.6)$$

$$R_{i,k} \triangleq \begin{cases} \frac{R_{k-d_i}}{l+1}, & k-d_i \leq N-d_l, \\ \frac{R_{k-d_i}}{j+1}, & N-d_{j+1} < k-d_i \leq N-d_j. \end{cases} \quad (3.7)$$

It is worthy mentioning that $\tilde{u}_s(k)$ is a combination of the past control inputs.

In order to derive a feedback control based on the current state, we consider the delay system (3.1) starting from an arbitrary given time $0 \leq s \leq N$ rather than 0.

Using the above notation, $\forall k \geq s$, the system (3.1) can be rewritten as

$$x(k+1) = \Phi(k)x(k) + \Gamma_s(k)u_s(k) + \tilde{u}_s(k). \quad (3.8)$$

The cost function (3.2) can be splitted into two terms,

$$J_N = J_N^s + \bar{J}_N^s,$$

where

$$J_N^s = \sum_{i=s}^N [u'_s(i)R_s(i)u_s(i) + x'(i)Q_i x(i)] + x'_{N+1}P_{N+1}x_{N+1}, \quad (3.9)$$

$$\bar{J}_N^s = \sum_{i=0}^{s-1} [u'(i)R_i u(i) + x'(i)Q_i x(i)]. \quad (3.10)$$

Note that $\{u(k), k \geq s\}$ has no effect on $\bar{J}_N^s$. So, we design $\{u(k), k \geq s\}$ to minimize J_N^s .

3.2.1 Duality between LQR and Smoothing

We shall introduce a dual backward stochastic system and establish a duality between the LQR problem for (3.1) and a smoothing problem for the backward stochastic system. Some elementary tools such as projection theory can be used for solving the smoothing problem for the backward stochastic system.

The stochastic backward system associated with system (3.8) and performance index (3.9) is given by

$$\mathbf{x}(k) = \Phi'(k)\mathbf{x}(k+1) + \mathbf{q}(k), \mathbf{x}(N+1) = \mathbf{x}_{N+1}, \quad (3.11)$$

$$\mathbf{y}(k) = \Gamma'_s(k)\mathbf{x}(k+1) + \mathbf{v}(k), s \leq k \leq N, \quad (3.12)$$

where $\mathbf{x}_{N+1}$, $\mathbf{q}(k)$ and $\mathbf{v}(k)$ are uncorrelated white noises with zero means and

covariances

$$\langle \mathbf{x}_{N+1}, \mathbf{x}_{N+1} \rangle = P_{N+1}, \quad \langle \mathbf{q}(i), \mathbf{q}(j) \rangle = Q_i \delta_{i,j}, \quad \langle \mathbf{v}(i), \mathbf{v}(j) \rangle = R_s(i) \delta_{i,j}.$$

By introducing the notation,

$$u_s \triangleq \text{col}\{u_s(s), \dots, u_s(N)\}, \quad \mathbf{y}_s \triangleq \text{col}\{\mathbf{y}(s), \dots, \mathbf{y}(N)\}, \quad h_s \triangleq (d_l - 1) \wedge (N - s), \quad (3.13)$$

we have the following result.

Lemma 3.1. *By making use of the stochastic state-space model (3.11)-(3.12), J_N^s can be put in the following quadratic form*

$$J_N^s = \begin{bmatrix} \xi_s \\ u_s \end{bmatrix}' \Pi_s \begin{bmatrix} \xi_s \\ u_s \end{bmatrix}, \quad (3.14)$$

where

$$\begin{aligned} \Pi_s &= \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ \mathbf{y}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ \mathbf{y}_s \end{bmatrix} \right\rangle \triangleq \begin{bmatrix} R_{\mathbf{x}_s^0} & R_{\mathbf{x}_s^0 \mathbf{y}_s} \\ R_{\mathbf{y}_s \mathbf{x}_s^0} & R_{\mathbf{y}_s} \end{bmatrix}, \\ \xi_s &\triangleq \left[x'(s) \tilde{u}'_s(s) \cdots \tilde{u}'_s(s + h_s) \right]', \quad \mathbf{x}_s^0 \triangleq \text{col}\{\mathbf{x}(s), \mathbf{x}(s+1), \dots, \mathbf{x}(s + h_s + 1)\}, \end{aligned}$$

with $\tilde{u}_s(s+i)$ as defined in (3.5), $\langle \mathbf{x}_s^0, \mathbf{x}_s^0 \rangle = R_{\mathbf{x}_s^0}$, $\langle \mathbf{x}_s^0, \mathbf{y}_s \rangle = R_{\mathbf{x}_s^0 \mathbf{y}_s}$ and $\langle \mathbf{y}_s, \mathbf{y}_s \rangle = R_{\mathbf{y}_s}$.

Proof. The proof is similar to that in [135] and [136] and is omitted. $\square$

Remark 3.2. *The backward stochastic system (3.11)-(3.12) is often used to convert the optimal control problem into an estimation problem [42], [135]. This is in fact motivated by the well known duality between the LQR for delay-free systems and the Kalman filtering. However, different from [135], we cannot directly solve the LQR problem based on this duality since it will lead to a constrained optimization problem (due to that $u_s(k)$ of (3.3) contains elements of the same input with different delays). We will derive a new backward stochastic system corresponding to system (3.1) and performance index (3.9) based on system (3.11)-(3.12).*

Due to the coupling of $u_s(k)$ at different delayed time instants, we apply a re-organization of measured outputs to decouple the u_s . First, decompose $\mathbf{y}(k)$ and $\mathbf{v}(k)$ of (3.12) as follows:

$$\mathbf{y}(k) \triangleq \begin{cases} \text{col}\{\mathbf{y}_0(k), \dots, \mathbf{y}_i(k)\}, d_i \leq k - s < d_{i+1}, \\ \text{col}\{\mathbf{y}_0(k), \dots, \mathbf{y}_l(k)\}, k - s \geq d_l, \end{cases} \quad (3.15)$$

$$\mathbf{v}(k) \triangleq \begin{cases} \text{col}\{\mathbf{v}_0(k), \dots, \mathbf{v}_i(k)\}, d_i \leq k - s < d_{i+1}, \\ \text{col}\{\mathbf{v}_0(k), \dots, \mathbf{v}_l(k)\}, k - s \geq d_l, \end{cases} \quad (3.16)$$

where $\mathbf{y}_i(k)$ and $\mathbf{v}_i(k)$, $i = 0, \dots, l$ satisfy

$$\mathbf{y}_i(k) = \Gamma'_{(i)}(k)\mathbf{x}(k+1) + \mathbf{v}_i(k), \quad (3.17)$$

with $\langle \mathbf{v}_i(k_1), \mathbf{v}_j(k_2) \rangle = R_{i,k_1} \delta_{i,j} \delta_{k_1,k_2}$. We can re-organize the outputs as follows:

$$\mathbf{y}_f(k) \triangleq \begin{cases} \text{col}\{\mathbf{y}_0(k), \dots, \mathbf{y}_l(k + d_l)\}, d_l \leq N - k, \\ \text{col}\{\mathbf{y}_0(k), \dots, \mathbf{y}_i(k + d_i)\}, d_i \leq N - k < d_{i+1}, \end{cases} \quad (3.18)$$

and correspondingly,

$$\mathbf{v}_f(k) \triangleq \begin{cases} \text{col}\{\mathbf{v}_0(k), \dots, \mathbf{v}_l(k + d_l)\}, d_l \leq N - k, \\ \text{col}\{\mathbf{v}_0(k), \dots, \mathbf{v}_i(k + d_i)\}, d_i \leq N - k < d_{i+1}. \end{cases} \quad (3.19)$$

Now, we shall simplify the cost function J_N^s of (3.14). To this end, we denote, for $s \leq k \leq N - d_l$,

$$T(k) = \underbrace{\begin{bmatrix} I_m & I_m & \cdots & I_m \\ 0 & -I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -I_m \end{bmatrix}}_{l+1}, \quad (3.20)$$

and for $N - d_{i+1} < k \leq N - d_i$,

$$T(k) = \underbrace{\begin{bmatrix} I_m & I_m & \cdots & I_m \\ 0 & -I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -I_m \end{bmatrix}}_{i+1}. \quad (3.21)$$

It is easy to write

$$T(k)\mathbf{y}_f(k) = \begin{cases} \text{col}\{\bar{\mathbf{y}}(k), -\mathbf{y}_1(k+d_1), \dots, -\mathbf{y}_l(k+d_l)\}, d_l \leq N-k, \\ \text{col}\{\bar{\mathbf{y}}(k), -\mathbf{y}_1(k+d_1), \dots, -\mathbf{y}_i(k+d_i)\}, d_i \leq N-k < d_{i+1}, \end{cases} \quad (3.22)$$

where

$$\bar{\mathbf{y}}(k) = \sum_{j=0}^l \mathbf{y}_j(k+d_j)\chi_{[k,N]}(k+d_j). \quad (3.23)$$

In view of (3.17), $\bar{\mathbf{y}}(k)$ can be rewritten as

$$\bar{\mathbf{y}}(k) = \sum_{j=0}^l \Gamma'_{(j)}(k+d_j)\mathbf{x}(k+d_j+1)\chi_{[k,N]}(k+d_j) + \bar{\mathbf{v}}(k), \quad (3.24)$$

where

$$\bar{\mathbf{v}}(k) = \sum_{j=0}^l \mathbf{v}_j(k+d_j)\chi_{[k,N]}(k+d_j). \quad (3.25)$$

It is easy to verify that $\bar{\mathbf{v}}(k), k = s, \dots, N$ are uncorrelated white noises with zero means and covariances $\langle \bar{\mathbf{v}}(i), \bar{\mathbf{v}}(k) \rangle = R_i \delta_{i,k}$. Then we have the following result.

Lemma 3.2. *Let*

$$\bar{\mathbf{y}}_s \triangleq \text{col}\{\bar{\mathbf{y}}(s), \dots, \bar{\mathbf{y}}(N)\}, \quad U_s \triangleq \text{col}\{u(s), \dots, u(N)\}, \quad (3.26)$$

where $\bar{\mathbf{y}}(k)$ is as in (3.24) and $u(k)$ in (3.1). Then J_N^s of (3.9) can be further

rewritten in the following quadratic form

$$J_N^s = \begin{bmatrix} \xi_s \\ U_s \end{bmatrix}' \bar{\Pi}_s \begin{bmatrix} \xi_s \\ U_s \end{bmatrix}, \quad (3.27)$$

where

$$\bar{\Pi}_s = \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ \bar{\mathbf{y}}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ \bar{\mathbf{y}}_s \end{bmatrix} \right\rangle \triangleq \begin{bmatrix} R_{\mathbf{x}_s^0} & R_{\mathbf{x}_s^0 \bar{\mathbf{y}}_s} \\ R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} & R_{\bar{\mathbf{y}}_s} \end{bmatrix}. \quad (3.28)$$

Furthermore, J_N^s can be simplified as

$$J_N^s = \xi_s' \mathcal{P}_s \xi_s + (U_s - U_s^*)' R_{\bar{\mathbf{y}}_s} (U_s - U_s^*), \quad (3.29)$$

where

$$U_s^* = -R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} \xi_s, \quad (3.30)$$

$$\mathcal{P}_s = \langle \mathbf{x}_s^0 - \hat{\mathbf{x}}_s^0, \mathbf{x}_s^0 - \hat{\mathbf{x}}_s^0 \rangle, \quad (3.31)$$

and $\hat{\mathbf{x}}_s^0$ is the projection of $\mathbf{x}_s^0$ onto the linear space $\mathcal{L}\{\bar{\mathbf{y}}_s\}$. Therefore, by considering (3.26), the minimizing solution of J_N^s with respect to control input $u(k)$ is the $(k - s + 1)$ -th block of U_s^* .

Proof. It can be proved similarly to [136]. The detail is thus omitted. $\square$

Remark 3.3. Lemma 3.2 reveals a duality between the LQR and the optimal smoothing problem for the stochastic backward system (3.11) and (3.24) in view of the fact that $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0}$ is the transpose of the optimal smoother gain of the backward system (3.11) and (3.24). However, it is not straightforward to derive an explicit optimal control based on the duality. Some further manipulations are needed, which are demonstrated in the next subsection.

3.2.2 Solution to the Finite Horizon LQR Problem

In view of the duality between the LQR problem and the smoothing problem, the controller gain matrix $-R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0}$ in (3.30) can be obtained from the following

projection problem:

$$\hat{\mathbf{x}}_s^0 = R_{\mathbf{x}_s^0 \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1} \bar{\mathbf{y}}_s. \quad (3.32)$$

So, we will derive the smoothing (filtering) gain matrix first. As in the standard Kalman filtering, we now define the cross covariance matrices of the state estimation error as follows:

$$P(k+i, k+j, k) \triangleq \langle \mathbf{x}(k+i), \tilde{\mathbf{x}}(k+j|k) \rangle, \quad (3.33)$$

Then the state estimation of the backward system (3.11) and (3.24) can be recursively calculated by the following theorem.

Theorem 3.1. *For system (3.11) and (3.24), the recursive optimal filter and smoother are given by*

$$\hat{\mathbf{x}}(k|k) = \Phi'(k) \hat{\mathbf{x}}(k+1|k), \quad (3.34)$$

$$\hat{\mathbf{x}}(k+i|k) = \hat{\mathbf{x}}(k+i|k+1) + K(k+i, k) \mathbf{e}(k), \quad i = 1, \dots, h_k + 1, \quad (3.35)$$

$$\mathbf{e}(k) = \bar{\mathbf{y}}(k) - \sum_{j=0}^l \Gamma'_{(j)}(k+d_j) \hat{\mathbf{x}}(k+d_j+1|k+1) \chi_{[k,N]}(k+d_j), \quad (3.36)$$

$$K(k+i, k) = \sum_{j=0}^l P(k+i, k+d_j+1, k+1) \Gamma_{(j)}(k+d_j) \chi_{[k,N]}(k+d_j) R_e^{-1}(k), \quad (3.37)$$

$$K(k, k) = \Phi'(k) K(k+1, k), \quad (3.38)$$

$$\begin{aligned} R_e(k) = & \sum_{i,j=0}^l [\Gamma'_{(i)}(k+d_i) P(k+d_i+1, k+d_j+1, k+1) \\ & \times \Gamma_{(j)}(k+d_j) \chi_{[k,N]}(k+d_i) \chi_{[k,N]}(k+d_j)] + R_k, \end{aligned} \quad (3.39)$$

where $P(k+i, k+j, k)$ is calculated by

$$\begin{aligned} P(k+i, k+j, k) = & P(k+i, k+j, k+1) - K(k+i, k) R_e(k) K'(k+j, k), \\ & i = 1, \dots, j, \quad j = 1, \dots, h_k + 1, \end{aligned} \quad (3.40)$$

$$P(k, k+j, k) = \Phi'(k) P(k+1, k+j, k), \quad j = 1, \dots, h_k + 1, \quad (3.41)$$

$$P(k+i, k+j, k) = P'(k+j, k+i, k), \quad (3.42)$$

$$P(k, k, k) = \Phi'(k)P(k+1, k+1, k)\Phi(k) + Q_k, \quad (3.43)$$

while h_k is defined in (3.13) with s replaced by k and $\mathbf{e}(k)$ is the innovation sequence with covariance $R_e(k)$.

The initial condition is:

$$\hat{\mathbf{x}}(N+1|N+1) = 0, \quad P(N+1, N+1, N+1) = P_{N+1}. \quad (3.44)$$

Proof. Applying the projection property, we can get

$$\hat{\mathbf{x}}(k+i|k) = \hat{\mathbf{x}}(k+i|k+1) + \langle \mathbf{x}(k+i), \mathbf{e}(k) \rangle R_e^{-1}(k) \mathbf{e}(k), \quad (3.45)$$

where $R_e(k) = \langle \mathbf{e}(k), \mathbf{e}(k) \rangle$ and

$$\mathbf{e}(k) = \sum_{j=0}^l \Gamma'_{(j)}(k+d_j) \tilde{\mathbf{x}}(k+d_j+1|k+1) \chi_{[k,N]}(k+d_j) + \bar{\mathbf{v}}(k). \quad (3.46)$$

Using the fact that $\forall i, j, k > 0$, $\langle \hat{\mathbf{x}}(i|k), \tilde{\mathbf{x}}(j|k) \rangle = 0$, $\langle \mathbf{x}(i), \bar{\mathbf{v}}(j) \rangle = 0$, we have

$$\begin{aligned} K(k+i, k) &= \langle \mathbf{x}(k+i), \mathbf{e}(k) \rangle R_e^{-1}(k) \\ &= \left\langle \tilde{\mathbf{x}}(k+i|k+1), \sum_{j=0}^l \Gamma'_{(j)}(k+d_j) \tilde{\mathbf{x}}(k+d_j+1|k+1) \chi_{[k,N]}(k+d_j) \right\rangle \\ &\quad \cdot R_e^{-1}(k) \\ &= \sum_{j=0}^l P(k+i, k+d_j+1, k+1) \Gamma_{(j)}(k+d_j) \chi_{[k,N]}(k+d_j) R_e^{-1}(k), \end{aligned}$$

which is (3.37). Substituting $K(k+i, k)$ into (3.45), we can get (3.35). Considering the system (3.11), it is easy to know that $\mathbf{q}(k)$ is orthogonal to $\mathcal{L}\{\bar{\mathbf{y}}(k), \dots, \bar{\mathbf{y}}(N)\}$, thereby (3.34) can easily be verified. (3.38) can be derived in the same way by considering (3.34). (3.39) can be verified by considering the expression of $\mathbf{e}(k)$ in (3.46). (3.42) is straightforward by virtue of the definition (3.33). According to

(3.45), we can write

$$\tilde{\mathbf{x}}(k+i|k) + K(k+i, k)\mathbf{e}(k) = \tilde{\mathbf{x}}(k+i|k+1). \quad (3.47)$$

Using the fact that $\mathbf{e}(k)$ is orthogonal to $\tilde{\mathbf{x}}(k+i|k)$, we can easily get

$$P(k+i, k+j, k) + K(k+i, k)R_e(k)K'(k+j, k) = P(k+i, k+j, k+1),$$

which is (3.40). Subtracting (3.34) from (3.11) yields

$$\tilde{\mathbf{x}}(k|k) = \Phi'(k)\tilde{\mathbf{x}}(k+1|k) + \mathbf{q}(k). \quad (3.48)$$

By multiplying the both sides of (3.48) with $\tilde{\mathbf{x}}'(k+j|k)$ from the right and taking the expectation, we arrive at (3.41). (3.43) is derived similarly by setting $j = 0$ in the last manipulation. $\square$

Theorem 3.1 provides a way to calculate $\tilde{\mathbf{x}}_s^0$, based on which we can obtain the result of the finite horizon LQR as follows.

Theorem 3.2. *Consider system (3.1) and cost function (3.2). The optimal LQR controller $u^*(s)$ that minimizes (3.2) is calculated by*

$$u^*(s) = -K'(s, s)x(s) - \sum_{i=0}^{h_s} K'(s+i+1, s)\tilde{u}_s^*(s+i), \quad (3.49)$$

where h_s is as defined in (3.13). $\tilde{u}_s^*(k)$ is given in (3.5) with $u(k-d_i)$ replaced by $u^*(k-d_i)$. $K(s+i, s), i = 0, \dots, h_s+1$ can be calculated by Theorem 3.1.

Furthermore, the minimum cost function is given by

$$J_N^* = x_0'P(0, 0, 0)x_0 + 2 \sum_{i=1}^{d_l} x_0'P(0, i, 0)\tilde{u}_0(i-1) + \sum_{i=1}^{d_l} \sum_{j=1}^{d_l} \tilde{u}_0'(i-1)P(i, j, 0)\tilde{u}_0(j-1). \quad (3.50)$$

Proof. From Theorem 3.1 we can write

$$\begin{aligned}
R_{\mathbf{x}_s^0, \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1} \bar{\mathbf{y}}_s &= \sum_{k=s}^N \langle \mathbf{x}_s^0, \mathbf{e}(k) \rangle R_e^{-1}(k) \mathbf{e}(k) \\
&= \langle \mathbf{x}_s^0, \mathbf{e}(s) \rangle R_e^{-1}(s) [\bar{\mathbf{y}}(s) - \hat{\bar{\mathbf{y}}}(s|s+1)] + \sum_{k=s+1}^N \langle \mathbf{x}_s^0, \mathbf{e}(k) \rangle R_e^{-1}(k) \mathbf{e}(k) \\
&= K_s \bar{\mathbf{y}}(s) + \bar{K}_s \bar{\mathbf{y}}_{s+1},
\end{aligned} \tag{3.51}$$

where

$$\begin{aligned}
K_s &= \text{col}\{K(s, s), K(s+1, s), \dots, K(s+h_s+1, s)\}, \\
\bar{K}_s \bar{\mathbf{y}}_{s+1} &= -K_s \hat{\bar{\mathbf{y}}}(s|s+1) + \sum_{k=s+1}^N \langle \mathbf{x}_s^0, \mathbf{e}(k) \rangle R_e^{-1}(k) \mathbf{e}(k), \\
\hat{\bar{\mathbf{y}}}(s|s+1) &= \sum_{j=0}^l \Gamma'_{(j)}(s+d_j) \hat{\mathbf{x}}(s+d_j+1|s+1) \chi_{[s, N]}(s+d_j).
\end{aligned}$$

The filter gain matrix $R_{\mathbf{x}_s^0, \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$ can be derived as

$$R_{\mathbf{x}_s^0, \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1} = \langle [K_s \ \bar{K}_s] \bar{\mathbf{y}}(s), \bar{\mathbf{y}}(s) \rangle R_{\bar{\mathbf{y}}_s}^{-1} = [K_s \ \bar{K}_s]. \tag{3.52}$$

In view of (3.30), the optimal controller can be given by

$$\begin{bmatrix} u^*(s) \\ \vdots \\ u^*(N) \end{bmatrix} = - \begin{bmatrix} K'_s \\ \bar{K}'_s \end{bmatrix} \begin{bmatrix} x(s) \\ \tilde{u}_s(s) \\ \vdots \\ \tilde{u}_s(s+h_s) \end{bmatrix}. \tag{3.53}$$

The first row block of (3.53) gives the optimal state feedback control law.

According to the definition of $\mathcal{P}_s$ in (3.31), the minimum cost is obtained as (3.50) by applying $U_s = U_s^*$ and setting $s = 0$ in (3.29). $\square$

Observe from (3.49) that the optimal control $u^*(k)$ is given in terms of the current state $x(k)$ and the previous input $u(k-i), i = 1, \dots, d_l$. Next, we will give a procedure for calculating the optimal controller gain matrices:

- 1) Set $s = N$, $P(N + 1, N + 1, N + 1) = P_{N+1}$;
- 2) Calculate $K(s + i, s)$, $i = 0, \dots, h_s + 1$ using (3.37), (3.38);
- 3) Update $P(s + i, s + j, s)$, $i, j = 0, \dots, h_s + 1$ using the relations (3.40)-(3.43);
- 4) $s \leftarrow s - 1$, if $s \geq 0$, goto 2); otherwise, end.

3.2.3 Solution to the Infinite Horizon LQR Problem

We shall study the LQR problem for linear discrete time-invariant systems in the infinite-horizon, i.e., we consider the system

$$x(k + 1) = \Phi x(k) + \sum_{i=0}^l \Gamma_{(i)} u(k - d_i), \quad x(0) = x_0, \quad (3.54)$$

with $u(k) = \mu_k$ when $-d_l \leq k < 0$, and the following quadratic performance index:

$$J_\infty = \lim_{N \rightarrow \infty} \sum_{i=0}^N [u'(i) R u(i) + x'(i) Q x(i)], \quad (3.55)$$

where $R > 0$, $Q \geq 0$ are constant weighting matrices on input signal u and state x , respectively.

The infinite horizon LQR problem is to find a state feedback control $\{u^*(s), s \geq 0\}$ such that the closed-loop system is asymptotically stable and the cost function J_∞ is minimized.

Different from the finite horizon case, for the infinite horizon case, we need to establish the stability of the closed-loop system. So, before extending the above result to the infinite horizon case, the analysis of the stability of the closed-loop system should be provided. To this end, we first denote

$$\bar{\Phi} = \begin{bmatrix} \Phi & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_n \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \bar{P}(k) = \begin{bmatrix} P(k, k, k) & \cdots & P(k, k + d_l, k) \\ \vdots & \ddots & \vdots \\ P(k + d_l, k, k) & \cdots & P(k + d_l, k + d_l, k) \end{bmatrix},$$

$$\bar{H} = [I_n \ 0 \ \cdots \ 0], \bar{\Gamma} = [\Gamma'_{(0)} \ \cdots \ \Gamma'_{(1)} \ \cdots \ \Gamma'_{(l)}]'$$

Using the above notation, we can write the Riccati recursions (3.40)-(3.43) in a compact form:

$$\bar{P}(k) = \bar{\Phi}'\bar{P}(k+1)\bar{\Phi} + \bar{H}'Q\bar{H} - K(k)R_e(k)K'(k), \quad (3.56)$$

where $K(k) = \bar{\Phi}'\bar{P}(k+1)\bar{\Gamma}R_e^{-1}(k)$, $R_e(k) = \bar{\Gamma}'\bar{P}(k+1)\bar{\Gamma} + R$.

Remark 3.4. Observe that the Riccati difference equation (3.56) and the Riccati recursions (3.40)-(3.43) are equivalent. In fact, if $P(k+i, k+j, k)$, $i, j = 0, \dots, h_k+1$ satisfy (3.40)-(3.43), then $\bar{P}(k)$ satisfies (3.56). Conversely, by direct calculation, when $\bar{P}(k)$ is the solution of (3.56), the sub-matrices of $\bar{P}(k)$ partitioned according to its definition will satisfy (3.40)-(3.43). Therefore, we can study the standard Riccati difference equation (3.56) instead of (3.40)-(3.43).

We recall the following lemma; see [9], [2] and [18].

Lemma 3.3. Consider the following ARE:

$$P = F'PF - F'PG(G'PG + R)^{-1}G'PF + Q, \quad (3.57)$$

where $Q \geq 0$ and $R > 0$. If (F, G) is stabilizable and $(Q^{\frac{1}{2}}, F)$ is detectable, then there exists a non-negative definite stabilizing solution $\bar{P}$ to the ARE (3.57), i.e. there exists a non-negative definite solution $\bar{P}$ such that the matrix $F - G(G'\bar{P}G + R)^{-1}G'\bar{P}F$ has all its eigenvalues inside the unit disk. Furthermore, for any given $P(0) \geq 0$, the solution of the Riccati difference equation

$$P(k+1) = F'P(k)F + Q - F'P(k)G[G'P(k)G + R]^{-1}G'P(k)F \quad (3.58)$$

exists and converges to the stabilizing solution $\bar{P}$, i.e. $\lim_{k \rightarrow \infty} P(k) = \bar{P}$.

Now, we are in the position to give the conditions guaranteeing the convergence of the Riccati recursions (3.40)-(3.43) and the stability of the closed-loop system.

Theorem 3.3. *For system (3.54) and cost function (3.55), if Φ is invertible, $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ is stabilizable and $(Q^{\frac{1}{2}}, \Phi)$ is detectable, then the solutions of Riccati recursions (3.40)-(3.43) converge and the closed-loop system of (3.54) with the converged control law given in Theorem 3.2 is asymptotically stable.*

Proof. Note that a backward Riccati difference equation has the same convergence property as its corresponding forward counterpart. We will first prove the equivalence between the stabilizability of $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ and that of $(\bar{\Phi}, \bar{\Gamma})$. Denote

$$[\lambda I - \bar{\Phi} \mid \bar{\Gamma}] = \left[\begin{array}{ccc|c} \lambda I - \Phi & -I & & \Gamma_{(0)} \\ & \lambda I & -I & \vdots \\ & & \ddots & \Gamma_{(1)} \\ & & & \lambda I & -I & \vdots \\ & & & & \lambda I & \Gamma_{(l)} \end{array} \right]. \quad (3.59)$$

If $(\bar{\Phi}, \bar{\Gamma})$ is not stabilizable, there exists a row vector $v = [v_0 \ v_1 \ \cdots \ v_d] \neq 0$ and a complex number λ , with $|\lambda| \geq 1$, such that $v [\lambda I - \bar{\Phi} \mid \bar{\Gamma}] = 0$, which implies

$$\left[v_0(\lambda I - \Phi) \quad \lambda v_1 - v_0 \quad \lambda v_2 - v_1 \quad \cdots \quad \sum_{i=0}^l v_{d_i} \Gamma_{(i)} \right] = 0. \quad (3.60)$$

From the above equation we see that

$$\lambda v_0 = v_0 \Phi, \quad v_i = v_0 \lambda^{-i}, \quad \sum_{i=0}^l v_{d_i} \Gamma_{(i)} = v_0 \sum_{i=0}^l \Gamma_{(i)} \lambda^{-d_i} = v_0 \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)} = 0.$$

Since $v \neq 0$, which implies $v_0 \neq 0$, it follows that $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ is not stabilizable, leading to a contradiction. The converse is also true. So the stabilizability of $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ and that of $(\bar{\Phi}, \bar{\Gamma})$ are equivalent. Similarly, we can prove that $(Q^{\frac{1}{2}} \bar{H}, \bar{\Phi})$ is detectable if and only if $(Q^{\frac{1}{2}}, \Phi)$ is detectable. On account of the assumption that $Q \geq 0$ and $R > 0$, if $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ is stabilizable and $(Q^{\frac{1}{2}}, \Phi)$ is detectable, the augmented Riccati difference equation (3.56) converges to the stabilizing solution according to Lemma 3.3. Hence, the Riccati difference recursions (3.40)-(3.43) converge.

From Theorem 3.2, the controller gain matrix is $-K'(k)$ in (3.56). The closed-loop system with the converged feedback gain can be written as

$$\xi_{k+1} = [\bar{\Phi} - \bar{\Gamma}(\bar{\Gamma}'\bar{P}\bar{\Gamma} + R)^{-1}\bar{\Gamma}'\bar{P}\bar{\Phi}] \xi_k. \quad (3.61)$$

Based on Lemma 3.3, we can conclude that system (3.61) is asymptotically stable. $\square$

Remark 3.5. *The condition that Φ is invertible is not very restrictive. In many practical applications, a discrete-time system often comes from discretization of a continuous system and the invertibility of Φ holds consequently. On the other hand, this assumption can be relaxed by replacing the stabilizability condition by that $(\Phi, \sum_{i=0}^l \lambda^{-d_i} \Gamma_{(i)})$ is stabilizable for all eigenvalues λ of Φ satisfying $|\lambda| \geq 1$.*

From Theorem 3.3 we know that under some standard condition the closed-loop system is stable, so we can design the optimal control using the result of finite horizon. The steady-state solution of the Riccati difference equation (3.40)-(3.43) is denoted as

$$\lim_{k \rightarrow \infty} P(k+i, k+j, k) = P(i, j), \quad i, j = 0, \dots, d_l,$$

and the corresponding $K(k+i, k), i = 0, \dots, d_l$ and $R_e(k)$ are denoted as $K(i), i = 0, \dots, d_l$ and R_e , respectively. Then, we have the following result.

Theorem 3.4. *Consider system (3.54) and cost function (3.55). If Φ is invertible, $(\Phi, \sum_{i=0}^l \Phi^{-d_i} \Gamma_{(i)})$ is stabilizable and $(Q^{\frac{1}{2}}, \Phi)$ is detectable, the infinite horizon LQR controller $u^*(s)$ exists and is given by*

$$u^*(s) = -K'(0)x(s) - \sum_{i=0}^{d_l-1} K'(i+1)\tilde{u}_s^*(s+i),$$

where $\tilde{u}_s^*(k)$ is obtained from (3.5) with $u(k-d_i)$ replaced by $u^*(k-d_i)$ and $K(i), i = 0, \dots, d_l$ can be calculated by

$$K(0) = \Phi'K(1), \quad K(i) = \sum_{j=0}^l P(i-1, d_j)\Gamma_{(j)}R_e^{-1}, \quad i = 1, \dots, d_l$$

with $R_e = \sum_{i,j=0}^l \Gamma'_{(i)} P(d_i, d_j) \Gamma_{(j)} + R$, and $P(i, j)$ satisfying the following Riccati recursions,

$$P(i, j) = P(i-1, j-1) - K(i)R_e K'(j), \quad i = 1, \dots, j, \quad j = 1, \dots, d_l, \quad (3.62)$$

$$P(0, j) = \Phi' P(1, j), \quad j = 1, \dots, d_l, \quad (3.63)$$

$$P(i, j) = P(j, i), \quad (3.64)$$

$$P(0, 0) = \Phi' P(1, 1) \Phi + Q. \quad (3.65)$$

Furthermore, the minimum cost function is given by

$$J_\infty^* = x_0' P(0, 0) x_0 + 2 \sum_{i=1}^{d_l} x_0' P(0, i) \tilde{u}_0(i-1) + \sum_{i=1}^{d_l} \sum_{j=1}^{d_l} \tilde{u}_0'(i-1) P(i, j) \tilde{u}_0(j-1).$$

Proof. The proof is straightforward by extending the finite horizon results and considering Theorem 3.3. $\square$

Remark 3.6. We note from Theorem 3.3 that under the standard stabilizability and detectability assumptions, the solutions to (3.62)-(3.65) can be computed by recursions.

3.3 Numerical Examples

Consider the system (3.1) as well as the cost function (3.2) with $l = 2$, $d_0 = 0$, $d_1 = 2$, $d_2 = 6$, $x_0 = [-1 \ 3]'$, $\mu_{-1} = \dots = \mu_{-6} = 0$, $N = 100$,

$$\Phi(k) = \begin{bmatrix} -3.141 & 5.875 \\ -1.99 & 2.511 \end{bmatrix}, \Gamma_0(k) = \begin{bmatrix} 1.815 \\ 0.927 \end{bmatrix}, \Gamma_1(k) = \begin{bmatrix} 0.906 \\ 0.913 \end{bmatrix}, \Gamma_2(k) = \begin{bmatrix} -1.688 \\ -0.076 \end{bmatrix},$$

$P_{N+1} = I_2$, $R_i = 1$, $Q_i = I_2$. Fig. 3.1 and Fig. 3.2 show the control signal and the two components of state in (3.1), respectively. The minimized cost function value is $J_N^* = 262.9248$.

Figure 3.1: State feedback control

3.4 Conclusion

This chapter has addressed the linear quadratic regulation problem for linear discrete-time systems with multiple delays in a single input channel. Compared with the work [134, 135], we presented a new and much simpler controller derivation approach. This approach enables us to derive an analytical solution to the LQR problem for systems with single input but multiple delays which cannot be addressed using the controller derivation method in [134, 135] since it will involve a constrained optimization problem. Through the new control design approach, an analytic optimal controller is given in terms of the solution of a two dimensional Riccati difference equation. The results were also extended to the infinite horizon case which has not been addressed in [134, 135]. Stability analysis has been provided.

Figure 3.2: Trajectory of state

Chapter 4

LQR for Continuous-time Systems with Multiple Delays in a Single Input Channel

In this chapter, we consider linear quadratic regulation for linear continuous-time systems with multiple delays in a single input channel. Different from the discrete-time counterpart, the augmentation approach cannot be applied. Using re-organization approach, we convert the LQR into smoothing problem for an associated backward stochastic system. The optimal feedback gain can be obtained by applying projection approach. A Riccati type partial differential equation should be solved for calculating the optimal controller. The result is also extended to infinite horizon case which is not trivial due to the stability problem for the closed-loop system. The result in this chapter is parallel to that in Chapter 3.

4.1 Problem Statement

Consider the following continuous-time linear system with multiple delays in the input channel

$$\dot{x}(t) = \Phi(t)x(t) + \sum_{i=0}^l \Gamma_{(i)}(t)u(t - d_i), x(0) = x_0, \quad (4.1)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input with initial value $u(t) = \nu_t$ when $-d_l \leq t < 0$, $\{d_i \in \bar{\mathbb{R}}^+, i = 0, \dots, l\}$ is the set of constant input time delays, $\Phi(t)$ and $\Gamma_{(i)}(t)$ are bounded time-varying matrices with appropriate dimension. Without loss of generality, the delays are also assumed to be of an increasing order. The following quadratic performance index will be introduced associated with the system (4.1):

$$J_T = x_T' P_T x_T + \int_0^T (u'(s) R_s u(s) + x'(s) Q_s x(s)) ds, \quad (4.2)$$

where $T > d_l$ is the time horizon, x_T is the terminal state, i.e., $x_T = x(T)$, $P_T \geq 0$ is the penalty weighting matrix on the terminal state, $\forall s \in [0, T]$, $R_s > 0$, $Q_s \geq 0$, are weighting matrices on the input signal $u(s)$ and the state $x(s)$, respectively.

The LQR problem for system (4.1) is to find the state feedback control $\{u^*(s), 0 \leq s \leq T\}$, such that the cost function J_T is minimized. When $T \rightarrow \infty$, the problem becomes the infinite horizon one.

Remark 4.1. *We note that the LQR problem for systems with single delay in each of the multiple input channels has been considered [134, 135]. It is not difficult to see that the class of systems considered in [134, 135] is a special case of (3.1) and (4.1) and the controller derivation method in [134, 135] cannot be readily extended for (3.1) and (4.1) as it will involve a constrained optimization. More specifically, recall that the class of the continuous-time systems considered in [134, 135] is of the form:*

$$\dot{x}(t) = \Phi(t)x(t) + \sum_{i=0}^l \tilde{\Gamma}_{(i)}(t)u_i(t - d_i). \quad (4.3)$$

By introducing

$$u(t) = [u'_1(t) \ u'_2(t) \ \cdots \ u'_l(t)]', \quad \Gamma_{(i)}(t) = [0 \ 0 \ \cdots \ \tilde{\Gamma}_{(i)}(t) \ 0 \ \cdots \ 0],$$

(4.3) is converted into (4.1). Same manipulation can be applied to the corresponding continuous-time systems. Therefore, a larger class of systems is considered in this chapter as compared to that of [134, 135]. Next, if we are to apply the optimal controller design approach proposed in [134, 135] to system (3.1) and (4.1), the following constraint

$$u_1(t) = u_2(t) = \cdots = u_l(t)$$

should be imposed, and the control problem becomes a constrained optimal control problem to which no analytical solution can be obtained.

4.2 Solution to the Continuous-time LQR Problem

Similar to the discrete-time system case in Chapter 3, we shall provide a parallel result of the LQR problem for continuous-time systems. By introducing the notation similar to (3.3)-(3.7), given $0 \leq s \leq T$, the system (4.1) can be rewritten as

$$\dot{x}(t) = \Phi(t)x(t) + \Gamma_s(t)u_s(t) + \tilde{u}_s(t), \quad \forall t \geq s, \quad (4.4)$$

and the cost function (4.2) can be splitted into two terms,

$$J_T = J_T^s + \bar{J}_T^s, \quad (4.5)$$

where

$$J_T^s = x'_T P_T x_T + \int_s^T (u'_s(\tau) R_s(\tau) u_s(\tau) + x'(\tau) Q_\tau x(\tau)) d\tau, \quad (4.6)$$

$$\bar{J}_T^s = \int_0^s (u'(\tau)R_\tau u(\tau) + x'(\tau)Q_\tau x(\tau)) d\tau. \quad (4.7)$$

$\{u(t), t \geq s\}$ can be designed to minimize only J_T^s because they have no effect on $\bar{J}_T^s$.

Remark 4.2. *The control problem for continuous-time systems with multiple delays in input channels has been studied in recent years, see [77], [134] for example. However, the LQR problem for systems with multiple delays in single input channel, which contains the problems addressed in [134] as a special case, is rather complicated. In [23], the problem is studied using an infinite dimensional systems theory approach which is difficult to understand. In what follows, we will provide a simple derivation and design method for the optimal control.*

4.2.1 Duality between LQR and Smoothing

Now, we introduce the following backward stochastic state-space system associated with (4.4) and performance index (4.6):

$$-\dot{\mathbf{x}}(t) = \Phi'(t)\mathbf{x}(t) + \mathbf{q}(t), \quad \mathbf{x}(T) = \mathbf{x}_T, \quad (4.8)$$

$$\mathbf{y}(t) = \Gamma'_s(t)\mathbf{x}(t) + \mathbf{v}(t), \quad s \leq t \leq T, \quad (4.9)$$

where $\mathbf{x}_T$, $\mathbf{q}(t)$ and $\mathbf{v}(t)$ are uncorrelated white noises with zero means and covariances

$$\langle \mathbf{x}_T, \mathbf{x}_T \rangle = P_T, \quad \langle \mathbf{q}(t_1), \mathbf{q}(t_2) \rangle = Q_{t_1} \delta(t_1 - t_2), \quad \langle \mathbf{v}(t_1), \mathbf{v}(t_2) \rangle = R_s(t_1) \delta(t_1 - t_2).$$

For a given $0 \leq s \leq T$, define

$$u_s \triangleq \{u_s(s + \tau), 0 \leq \tau \leq T - s\}, \quad \tilde{u}_s \triangleq \{\tilde{u}_s(s + \tau), 0 \leq \tau \leq h_s\}, \quad h_s \triangleq d_l \wedge (T - s). \quad (4.10)$$

Motivated by the discrete-time setting, $\forall A, B \in L_2\{a, b; \mathbb{R}^r\}$, we slightly abuse notation here by defining

$$A'B \triangleq \int_a^b A'(\tau)B(\tau)d\tau, \quad A'MA \triangleq \int_a^b \int_a^b A'(\tau_1)M(\tau_1, \tau_2)A(\tau_2)d\tau_1d\tau_2.$$

Then we have the following result.

Lemma 4.1. *By making use of the stochastic state-space model (4.8)-(4.9), J_T^s can be put in the following quadratic form*

$$J_T^s = \begin{bmatrix} \xi_s \\ u_s \end{bmatrix}' \Pi_s \begin{bmatrix} \xi_s \\ u_s \end{bmatrix}, \quad (4.11)$$

where

$$\begin{aligned} \Pi_s &= \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ \mathbf{y}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ \mathbf{y}_s \end{bmatrix} \right\rangle \triangleq \begin{bmatrix} R_{\mathbf{x}_s^0} & R_{\mathbf{x}_s^0 \mathbf{y}_s} \\ R_{\mathbf{y}_s \mathbf{x}_s^0} & R_{\mathbf{y}_s} \end{bmatrix}, \quad \xi_s \triangleq [x'(s) \quad \tilde{u}'_s]', \quad \mathbf{x}_s^0 \triangleq [\mathbf{x}(s)' \quad \mathbf{x}'_s]', \\ \mathbf{x}_s &\triangleq \{\mathbf{x}(s+\tau), 0 \leq \tau \leq h_s\}, \quad \mathbf{y}_s \triangleq \{\mathbf{y}(s+\tau), 0 \leq \tau \leq T-s\}, \end{aligned}$$

with $\langle \mathbf{x}_s^0, \mathbf{x}_s^0 \rangle = R_{\mathbf{x}_s^0}$, $\langle \mathbf{x}_s^0, \mathbf{y}_s \rangle = R_{\mathbf{x}_s^0 \mathbf{y}_s}$ and $\langle \mathbf{y}_s, \mathbf{y}_s \rangle = R_{\mathbf{y}_s}$.

Proof. Similar proof can be found in [135]. □

We reorganize the measured outputs by first decomposing $\mathbf{y}(t)$ and $\mathbf{v}(t)$ of (4.9) as follows:

$$\begin{aligned} \mathbf{y}(t) &\triangleq \begin{cases} \text{col} \{\mathbf{y}_0(t) \cdots \mathbf{y}_i(t)\}, & d_i \leq t-s < d_{i+1}, \\ \text{col} \{\mathbf{y}_0(t) \cdots \mathbf{y}_l(t)\}, & t-s \geq d_l, \end{cases} \\ \mathbf{v}(t) &\triangleq \begin{cases} \text{col} \{\mathbf{v}_0(t) \cdots \mathbf{v}_i(t)\}, & d_i \leq t-s < d_{i+1}, \\ \text{col} \{\mathbf{v}_0(t) \cdots \mathbf{v}_l(t)\}, & t-s \geq d_l, \end{cases} \end{aligned}$$

where $\mathbf{y}_i(t)$ and $\mathbf{v}_i(t)$, $i = 0, \dots, l$ satisfy

$$\mathbf{y}_i(t) = \Gamma'_{(i)}(t)\mathbf{x}(t) + \mathbf{v}_i(t), \quad (4.12)$$

with $\langle \mathbf{v}_i(t_1), \mathbf{v}_j(t_2) \rangle = R_{i,t_1} \delta_{i,j} \delta(t_1 - t_2)$. The outputs can be reorganized as follows:

$$\mathbf{y}_f(t) \triangleq \begin{cases} \text{col}\{\mathbf{y}_0(t), \dots, \mathbf{y}_l(t + d_l)\}, d_l \leq N - t, \\ \text{col}\{\mathbf{y}_0(t), \dots, \mathbf{y}_i(t + d_i)\}, d_i \leq N - t < d_{i+1}, \end{cases}$$

and $\mathbf{v}_f(t)$ has the same form as

$$\mathbf{v}_f(t) \triangleq \begin{cases} \text{col}\{\mathbf{v}_0(t), \dots, \mathbf{v}_l(t + d_l)\}, d_l \leq N - t, \\ \text{col}\{\mathbf{v}_0(t), \dots, \mathbf{v}_i(t + d_i)\}, d_i \leq N - t < d_{i+1}. \end{cases}$$

In order to simplify the cost function J_T^s of (4.11), by introducing $T(t)$ defined similarly in (3.20)-(3.21), we write

$$T(t)\mathbf{y}_f(t) = \begin{cases} \text{col}\{\bar{\mathbf{y}}(t), -\mathbf{y}_1(t + d_1), \dots, -\mathbf{y}_l(t + d_l)\}, d_l \leq T - t, \\ \text{col}\{\bar{\mathbf{y}}(t), -\mathbf{y}_1(t + d_1), \dots, -\mathbf{y}_i(t + d_i)\}, d_i \leq T - t < d_{i+1}, \end{cases}$$

where

$$\bar{\mathbf{y}}(t) = \sum_{j=0}^l \mathbf{y}_j(t + d_j) \chi_{[t,T]}(t + d_j).$$

In view of (4.12), $\bar{\mathbf{y}}(t)$ can be rewritten as

$$\bar{\mathbf{y}}(t) = \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \mathbf{x}(t + d_j) \chi_{[t,T]}(t + d_j) + \bar{\mathbf{v}}(t), \quad (4.13)$$

where

$$\bar{\mathbf{v}}(t) = \sum_{j=0}^l \mathbf{v}_j(t + d_j) \chi_{[t,T]}(t + d_j).$$

It is easy to verify that $\bar{\mathbf{v}}(t), s \leq t \leq T$ is a white noise process with zero means and covariances $\langle \bar{\mathbf{v}}(t_1), \bar{\mathbf{v}}(t_2) \rangle = R_{t_1} \delta(t_1 - t_2)$, where R_{t_1} appears in the cost function (4.2). Then we have the following result.

Lemma 4.2. *Let*

$$\bar{\mathbf{y}}_s \triangleq \{\bar{\mathbf{y}}(s + \tau), 0 \leq \tau \leq T - s\}, \quad U_s \triangleq \{u(s + \tau), 0 \leq \tau \leq T - s\}, \quad (4.14)$$

then J_T^s of (4.6) can be further rewritten in the following quadratic form

$$J_T^s = \begin{bmatrix} \xi_s \\ U_s \end{bmatrix}' \bar{\Pi}_s \begin{bmatrix} \xi_s \\ U_s \end{bmatrix}, \quad (4.15)$$

where

$$\bar{\Pi}_s = \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ \bar{\mathbf{y}}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ \bar{\mathbf{y}}_s \end{bmatrix} \right\rangle \triangleq \begin{bmatrix} R_{\mathbf{x}_s^0} & R_{\mathbf{x}_s^0 \bar{\mathbf{y}}_s} \\ R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} & R_{\bar{\mathbf{y}}_s} \end{bmatrix}.$$

Furthermore, J_T^s can be simplified as

$$J_T^s = \xi_s' \mathcal{P}_s \xi_s + (U_s - U_s^*)' R_{\bar{\mathbf{y}}_s} (U_s - U_s^*), \quad (4.16)$$

where

$$U_s^* = -R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} \xi_s, \quad (4.17)$$

$$\mathcal{P}_s = \langle \mathbf{x}_s^0 - \hat{\mathbf{x}}_s^0, \mathbf{x}_s^0 - \hat{\mathbf{x}}_s^0 \rangle, \quad (4.18)$$

and $\hat{\mathbf{x}}_s^0$ is the projection of $\mathbf{x}_s^0$ onto the linear space $\mathcal{L}\{\bar{\mathbf{y}}_s\}$.

Remark 4.3. The duality between the LQR and the optimal smoothing problem for the stochastic backward system (4.8) and (4.13) is established in view of the fact that $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0}$ is the transpose of the optimal smoothing gain of the backward system, which allows us to solve the LQR problem by projection theory [24].

4.2.2 Solution to the Finite Horizon LQR Problem

The controller gain $-R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_0^s}$ in (4.17) can be obtained from the following projection problem:

$$\hat{\mathbf{x}}_s^0 = R_{\mathbf{x}_s^0 \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1} \bar{\mathbf{y}}_s. \quad (4.19)$$

Next we will give a way to calculate the estimation of the state for backward system (4.8) and (4.13). Similar result can be found in [61].

Theorem 4.1. *For system (4.8) and (4.13), the recursive optimal filter and smoother are given by*

$$\dot{\hat{\mathbf{x}}}(t|t) = -\Phi'(t)\hat{\mathbf{x}}(t|t) - K(t, t)\mathbf{e}(t), \quad (4.20)$$

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right] \hat{\mathbf{x}}(t - \theta|t) = -K(t - \theta, t)\mathbf{e}(t), \quad -h_t \leq \theta \leq 0, \quad (4.21)$$

$$K(t - \theta, t) = \sum_{j=0}^l P(t - \theta, t + d_j, t) \Gamma_{(j)}(t + d_j) \chi_{[t, T]}(t + d_j) R_t^{-1}, \quad (4.22)$$

$$\mathbf{e}(t) = \bar{\mathbf{y}}(t) - \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \hat{\mathbf{x}}(t + d_j|k) \chi_{[t, T]}(t + d_j), \quad (4.23)$$

where h_t is defined in (4.10) with s replaced by t , $\mathbf{e}(t)$ is the innovation process with covariance R_t . The covariance matrix $P(\cdot, \cdot, \cdot)$ satisfies $\forall \theta, \theta_1, \theta_2 \in [-h_t, 0]$

$$\dot{P}(t, t, t) = -\Phi'(t)P(t, t, t) - P(t, t, t)\Phi(t) - Q_t + K(t, t)R_tK'(t, t), \quad (4.24)$$

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right] P(t - \theta, t, t) = -P(t - \theta, t, t)\Phi(t) + K(t - \theta, t)R_tK'(t, t), \quad (4.25)$$

$$P(t - \theta_1, t - \theta_2, t) = P'(t - \theta_2, t - \theta_1, t), \quad (4.26)$$

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta_1} + \frac{\partial}{\partial \theta_2}\right] P(t - \theta_1, t - \theta_2, t) = K(t - \theta_1, t)R_tK'(t - \theta_2, t). \quad (4.27)$$

The initial condition is:

$$\hat{\mathbf{x}}(T|T) = 0, \quad P(T, T, T) = P_T.$$

Proof. Applying the projection property, we can get

$$\hat{\mathbf{x}}(t|t) = \int_t^T \langle \mathbf{x}(t), \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau. \quad (4.28)$$

First we define $K(t - \theta, t) = \langle \mathbf{x}(t - \theta), \mathbf{e}(t) \rangle R_t^{-1}$. Take derivative of (4.28) and consider (4.8),

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(t|t) &= \int_t^T \langle \dot{\mathbf{x}}(t), \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau - \langle \mathbf{x}(t), \mathbf{e}(t) \rangle R_t^{-1} \mathbf{e}(t) \\ &= - \int_t^T \langle \Phi'(t)\mathbf{x}(t) + \mathbf{q}(t), \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau - K(t, t)\mathbf{e}(t). \end{aligned}$$

Noting the fact that $\forall \tau \geq t, \langle \mathbf{q}(t), \mathbf{e}(\tau) \rangle = 0$, (4.20) is obtained.

Similar to (4.28), it follows that

$$\hat{\mathbf{x}}(t - \theta|t) = \int_t^T \langle \mathbf{x}(t - \theta), \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau. \quad (4.29)$$

By taking partial derivative of (4.29) with respect to t and θ respectively and adding the results yield (4.21), $\mathbf{e}(t)$ can be written as

$$\mathbf{e}(t) = \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \tilde{\mathbf{x}}(t + d_j|k) \chi_{[t,T]}(t + d_j) + \bar{\mathbf{v}}(t). \quad (4.30)$$

Then $K(t - \theta, t), -h_t \leq \theta \leq 0$ can be expressed in the following form

$$K(t - \theta, t) = \left\langle \hat{\mathbf{x}}(t - \theta|t) + \tilde{\mathbf{x}}(t - \theta|t), \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \tilde{\mathbf{x}}(t + d_j|t) \chi_{[t,T]}(t + d_j) + \bar{\mathbf{v}}(t) \right\rangle \cdot R_t^{-1}.$$

Using the fact that $\forall \theta, \theta_1, \theta_2, \langle \mathbf{x}(t - \theta), \bar{\mathbf{v}}(t) \rangle = 0, \langle \hat{\mathbf{x}}(t - \theta_1|t), \tilde{\mathbf{x}}(t - \theta_2|t) \rangle = 0$, (4.22) is easy to be verified.

What remain to be derived are the differential equations (4.24)-(4.27). Adding (4.8) and (4.20) yields

$$\dot{\hat{\mathbf{x}}}(t|t) = -\Phi'(t) \tilde{\mathbf{x}}(t|t) - \mathbf{q}(t) + K(t, t) \mathbf{e}(t). \quad (4.31)$$

Then we write $\tilde{\mathbf{x}}(t|t)$ in the following form

$$\tilde{\mathbf{x}}(t|t) + \int_t^T \Psi_{t,\tau}^{-\Phi'} K(\tau, \tau) \mathbf{e}(\tau) d\tau = \Psi_{t,T}^{-\Phi'} \mathbf{x}_T + \int_t^T \Psi_{t,\tau}^{-\Phi'} \mathbf{q}(\tau) d\tau, \quad (4.32)$$

where $\Psi_{t,T}^{-\Phi'}$ is the transition matrix from T to t generated by $-\Phi'$. Note that all the terms on both sides of (4.32) are uncorrelated. Meanwhile, $P(t - \theta_1, t - \theta_2, t)$ can be rewritten as

$$P(t - \theta_1, t - \theta_2, t) = \langle \hat{\mathbf{x}}(t - \theta_1|t) + \tilde{\mathbf{x}}(t - \theta_1|t), \tilde{\mathbf{x}}(t - \theta_2|t) \rangle$$

$$= \langle \tilde{\mathbf{x}}(t - \theta_1 | t), \tilde{\mathbf{x}}(t - \theta_2 | t) \rangle. \quad (4.33)$$

By calculating the covariances of both sides of (4.32), we can get

$$\begin{aligned} P(t, t, t) &= \Psi_{t,T}^{-\Phi'} P_T \Psi_{T,t}^{-\Phi} + \int_t^T \int_t^T \Psi_{t,\tau_1}^{-\Phi'} Q_{\tau_1} \delta(\tau_1 - \tau_2) \Psi_{\tau_2,t}^{-\Phi} d\tau_1 d\tau_2 \\ &\quad - \int_t^T \int_t^T \Psi_{t,\tau_1}^{-\Phi'} K(\tau_1, \tau_1) R_{\tau_1} \delta(\tau_1 - \tau_2) K'(\tau_2, \tau_2) \Psi_{\tau_2,t}^{-\Phi} d\tau_1 d\tau_2 \\ &= \Psi_{t,T}^{-\Phi'} P_T \Psi_{T,t}^{-\Phi} + \int_t^T \Psi_{t,\tau}^{-\Phi'} Q_{\tau} \Psi_{\tau,t}^{-\Phi} d\tau - \int_t^T \Psi_{t,\tau}^{-\Phi'} K(\tau, \tau) R_{\tau} K'(\tau, \tau) \Psi_{\tau,t}^{-\Phi} d\tau. \end{aligned} \quad (4.34)$$

From the above equation, we see that $P(t, t, t)$ is differentiable. (4.24) is obtained by directly differentiating both sides of (4.34) with respect of t . Next, we shall investigate the differentiability of $P(t - \theta_1, t - \theta_2, t)$.

$$\begin{aligned} \frac{\partial P(t - \theta_1, t - \theta_2, t)}{\partial t} &\triangleq \lim_{\Delta \rightarrow 0} \frac{P(t + \Delta - \theta_1, t + \Delta - \theta_2, t + \Delta) - P(t - \theta_1, t - \theta_2, t)}{\Delta} \\ &= \lim_{\Delta \rightarrow 0} \left\langle \tilde{\mathbf{x}}(t + \Delta - \theta_1 | t + \Delta), \frac{\tilde{\mathbf{x}}(t + \Delta - \theta_2 | t + \Delta) - \tilde{\mathbf{x}}(t - \theta_2 | t)}{\Delta} \right\rangle \\ &\quad + \lim_{\Delta \rightarrow 0} \left\langle \frac{\tilde{\mathbf{x}}(t + \Delta - \theta_1 | t + \Delta) - \tilde{\mathbf{x}}(t - \theta_1 | t)}{\Delta}, \tilde{\mathbf{x}}(t - \theta_2 | t) \right\rangle \\ &= \left\langle \tilde{\mathbf{x}}(t - \theta_1 | t), \frac{\partial \tilde{\mathbf{x}}(t - \theta_2 | t)}{\partial t} \right\rangle + \left\langle \frac{\partial \tilde{\mathbf{x}}(t - \theta_1 | t)}{\partial t}, \tilde{\mathbf{x}}(t - \theta_2 | t) \right\rangle. \end{aligned}$$

From the above equation, we see that $P(t - \theta_1, t - \theta_2, t)$ is differentiable with respect to t . By the same argument, it is evident that $P(t - \theta_1, t - \theta_2, t)$ is differentiable with respect to both θ_1 and θ_2 . Therefore, we can write

$$\frac{\partial P(t - \theta, t, t)}{\partial t} = \left\langle \frac{\partial \mathbf{x}(t - \theta)}{\partial t}, \tilde{\mathbf{x}}(t | t) \right\rangle + \left\langle \mathbf{x}(t - \theta), \dot{\tilde{\mathbf{x}}}(t | t) \right\rangle, \quad (4.35)$$

$$\frac{\partial P(t - \theta, t, t)}{\partial \theta} = \left\langle \frac{\partial \mathbf{x}(t - \theta)}{\partial \theta}, \tilde{\mathbf{x}}(t | t) \right\rangle. \quad (4.36)$$

Adding (4.35) and (4.36) and considering (4.31) yield

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} \right] P(t - \theta, t, t) = \left\langle \mathbf{x}(t - \theta), -\Phi'(t) \tilde{\mathbf{x}}(t | t) - \mathbf{q}(t) + K(t, t) \mathbf{e}(t) \right\rangle$$

$$= -P(t - \theta, t, t)\Phi(t) + K(t - \theta, t)R_t K'(t, t),$$

which is (4.25). (4.27) can be proved analogously and (4.26) holds by considering (4.33). $\square$

Theorem 4.1 provides a way to calculate $\hat{\mathbf{x}}_s^0$, based on which we can obtain the result of the finite horizon LQR as follows.

Theorem 4.2. *Consider system (4.1) and cost function (4.2). The optimal LQR controller $u^*(s)$ that minimizes (4.2) is calculated by*

$$u^*(s) = -K'(s, s)x(s) - \int_{-h_s}^0 K'(s - \tau, s)\tilde{u}_s^*(s - \tau)d\tau, \quad (4.37)$$

where h_s is as defined in (4.10). $\tilde{u}_s^*(t)$ is given by replacing $u^*(t - d_i)$ for $u(t - d_i)$ in $\tilde{u}_s(t)$. $K(s - \tau, s)$, $-h_s \leq \tau \leq 0$ can be calculated by Theorem 4.1.

Furthermore, the minimum value of J_T is given by

$$\begin{aligned} J_T^* &= x_0' P(0, 0, 0)x_0 + 2 \int_0^{d_1} x_0' P(0, \theta, 0)\tilde{u}_0(\theta)d\theta \\ &\quad + \int_0^{d_1} \int_0^{d_1} \tilde{u}_0(\theta_1)P(\theta_1, \theta_2, 0)\tilde{u}_0(\theta_2)d\theta_1 d\theta_2. \end{aligned}$$

Proof. Define a linear operator $\mathcal{G}$ by $\mathbf{e}_s = \mathcal{G}\bar{\mathbf{y}}_s$, where $\mathbf{e}_s = \{\mathbf{e}(s + \tau), 0 \leq \tau \leq T - s\}$. From the proof of Theorem 4.1, we write (4.28) and (4.29) in the following form

$$\hat{\mathbf{x}}_s^0 = \int_s^T \langle \mathbf{x}_s^0, \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau \triangleq \mathcal{S} \mathbf{e}_s,$$

where $\mathcal{S}$ is a linear operator mapping $\mathbf{x}_s^0$ on $\mathcal{L}\{\mathbf{e}_s\}$. In view of (4.17), the optimal controller can be given by

$$U_s^* = -\mathcal{G}^* \mathcal{S}^* \xi_s,$$

where $\mathcal{S}^*$ is the adjoint operator of $\mathcal{S}$. By introducing Dirac distribution function T_{δ_r} , which is a generalized function concentrated at the point r , the following equation

is easy to be verified

$$u(s) = -T_{\delta_0} \mathcal{G}^* \mathcal{S}^* \xi_s = -K'(s, s)x(s) - \int_{-h_s}^0 K'(s - \tau, s) \tilde{u}_s^*(s - \tau) d\tau,$$

which is (4.37).

By substituting $U_s = U_s^*$ and $s = 0$ into (4.16) and considering (4.18), the minimum cost function is obtained easily. $\square$

Remark 4.4. *Theorem 4.2 presents a solution to the LQR problem for linear continuous-time systems with multiple delays in a single input channel. The solution is given in terms of a Riccati type partial differential equation (4.24)-(4.27) which can be solved using some numerical methods such as [37].*

Remark 4.5. *In existing literature, a popular way of solving the control problem for continuous systems with delay in controller is to convert the delay systems into delay-free distributed parameter systems [94], [56]. The controller is given in terms of the solution of an operator Riccati equation which generally has no explicit solution except for some particular cases such as those in [58], [112]. The proposed approach can directly give an explicit solution without the need of converting the operator Riccati equation into a finite dimensional one. We will present a new way for constructing an infinite dimensional system. We extend $\tilde{u}$ into the state rather than directly extend the past inputs into the state. This technique is similar to that in [23] but is originated from a rather different method of derivation.*

Remark 4.6. *For systems with single input delay or multiple input channel with single delay in each channel, the controller can be given in terms of the solution of a Riccati differential equation. However, for the systems under investigation, PDE is unavoidable due to the fact that the current input affects not only the current state but also future state.*

Next, we shall investigate the convergence property of the Riccati partial differential equation (4.24)-(4.27) associated with the LQR solution and the stability of the

closed-loop system by infinite dimensional system theory. This enables us to extend the above results to the infinite horizon case.

4.2.3 Solution to the Infinite Horizon LQR Problem

We shall study the LQR problem for linear continuous time-invariant systems in the infinite-horizon, i.e., we consider the system

$$\dot{x}(t) = \Phi x(t) + \sum_{i=0}^l \Gamma_{(i)} u(t - d_i), x(0) = x_0, \quad (4.38)$$

with $u(t) = \nu_t$ when $-d_l \leq t < 0$, and the following quadratic performance index:

$$J_\infty = \lim_{T \rightarrow \infty} \int_0^T (u'(\tau) R u(\tau) + x'(\tau) Q x(\tau)) d\tau, \quad (4.39)$$

where $R > 0$, $Q \geq 0$ are weighting matrices on input signal u and state x , respectively.

The infinite horizon LQR problem is to find a state feedback control $u^*(s)$, such that the closed-loop system is asymptotically stable and the cost function J_∞ is minimized.

As before, the analysis of the stability of the closed-loop system should be given first. We will reformulate the system (4.38) using an abstract evolution model. We introduce the well-known Hilbert space $\mathcal{M}_2$ as $\mathbb{R}^n \times L_2(-d_l, 0; \mathbb{R}^n)$ with the obvious inner product, and the input space $\mathcal{U}$ as $\mathbb{R}^m$ endowed with standard inner product. Then the abstract evolution realization of the system (4.38) will be formally written as

$$\frac{d\phi(s)}{ds} = \mathcal{A}\phi(s) + \mathcal{B}u(s), \phi(0) = (x_0, \tilde{u}_0) \in \mathcal{M}_2, \quad (4.40)$$

where $\tilde{u}_0$ is as defined in (4.10), $\phi(s) = (f_s^0, f_s^1(\tau))$ with

$$f_s^0 = x(s), \quad f_s^1(\tau) = \tilde{u}_s(s - \tau), \quad \tau \in [-d_l, 0].$$

Denote by D the differential operator, i.e. $Dy(\tau) = \frac{dy(\tau)}{d\tau}$. The coefficients in (4.40) are defined as follows:

$$\mathcal{A}(f, g) = (\Phi f + g(0), -Dg), \quad (4.41)$$

$$\mathcal{B}u = \left(\Gamma_{(0)}u, \sum_{i=1}^l \delta(\tau + d_i) \Gamma_{(i)}u \right), \quad \tau \in [-d_l, 0]. \quad (4.42)$$

It will be established that $\mathcal{A}$ generates a C_0 -semigroup on $\mathcal{M}_2$ with domain

$$\mathcal{D}(\mathcal{A}) = \{(f, g) \in \mathcal{M}_2 \mid g \in W^1(-d_l, 0; \mathbb{R}^n), g(-d_l) = 0\}.$$

Remark 4.7. *It is obvious that the input operator $\mathcal{B}$ is an unbounded operator from $\mathcal{U}$ to $\mathcal{M}_2$ due to the existence of δ . However, $\mathcal{B}$ is indeed a linear bounded operator from $\mathcal{U}$ to a ‘larger’ space $\mathcal{D}(\mathcal{A}^*)^*$, or equivalently, there exists a $\lambda_0 \in \rho(\mathcal{A})$ such that $\mathcal{B}$ is of the form $\mathcal{B} = (\lambda_0 \mathcal{I} - \mathcal{A}) \mathcal{J}$, where $\mathcal{J}$ is a linear bounded operator from $\mathcal{U}$ to $\mathcal{M}_2$. The differential equation (4.40) for the state still makes sense for this control operator. An explanation will be given later.*

We will characterize the adjoint operator $\mathcal{A}^*$ of $\mathcal{A}$ in the following theorem.

Theorem 4.3. *$\mathcal{A}^*$ is an infinitesimal generator on $\mathcal{M}_2$ with domain*

$$\mathcal{D}(\mathcal{A}^*) = \{(z, w) \in \mathcal{M}_2 \mid w \in W^1(-d_l, 0; \mathbb{R}^n), w(0) = z\},$$

and defined by $\mathcal{A}^(z, w) = (\Phi'z, Dw)$.*

Proof. The proof is similar to that in [50] and thus omitted. $\square$

Analogously, it is easy to verify that $\mathcal{B}^*$ (the adjoint operator of $\mathcal{B}$) is defined as

$$\mathcal{B}^*(z, w) = \sum_{i=0}^l \Gamma'_{(i)} w(-d_i). \quad (4.43)$$

$\mathcal{B}^*$ is also an unbounded operator from $\mathcal{M}_2$ to $\mathcal{U}$. However, it is a linear bounded operator from $\mathcal{D}(\mathcal{A}^*)$ (dense in $\mathcal{M}_2$) to $\mathcal{U}$. In fact, the evolution realization can be

regarded as a smooth Pritchard-Salamon system by defining the space $\mathcal{V} = \mathcal{D}(\mathcal{A}^*)^*$ in [56]. Next, we shall find out the exact expression of $\mathcal{J}$.

Lemma 4.3. *For linear operator $\mathcal{A}$ and $\mathcal{B}$ defined in (4.41) and (4.42), respectively, $\exists \lambda_0 \in \rho(\mathcal{A})$ such that*

$$\mathcal{B} = (\lambda_0 \mathcal{I} - \mathcal{A}) \mathcal{J}, \quad (4.44)$$

where $\mathcal{I}$ is the identity operator, $\mathcal{J}$ is a bounded linear operator from $\mathcal{U}$ to $\mathcal{M}_2$.

Proof. First, we will show that the spectrum of $\mathcal{A}$ is a pure point spectrum and $\sigma(\mathcal{A}) = \sigma(\Phi)$. According to [56], we assume that there exists a λ and $(z, w) \neq 0 \in \mathcal{D}(\mathcal{A})$ such that

$$(\lambda \mathcal{I} - \mathcal{A})(z, w) = 0. \quad (4.45)$$

Then we have

$$\lambda z = \Phi z + w(0), \quad \lambda w(\theta) + Dw(\theta) = 0, \quad w(-d_l) = 0. \quad (4.46)$$

The solution of (4.46) is $w(\theta) = 0$. It implies that the point spectrum of $\mathcal{A}$ is equal to $\sigma(\Phi)$. Next, let $\lambda \notin \sigma(\Phi)$, we shall show that $\lambda \in \rho(\mathcal{A})$. Suppose that $(\lambda \mathcal{I} - \mathcal{A})(x, y) = (z, w)$. It follows that $(x, y) = (\lambda \mathcal{I} - \mathcal{A})^{-1}(z, w)$ is given by

$$x = (\lambda I - \Phi)^{-1} \left(z + \int_{-d_l}^0 e^{\lambda \tau} w(\tau) d\tau \right), \quad y(\theta) = \int_{-d_l}^{\theta} e^{\lambda(\tau-\theta)} w(\tau) d\tau.$$

From the above we see that $(\lambda \mathcal{I} - \mathcal{A})^{-1}$ is a linear operator on $\mathcal{M}_2$, which implies that $\lambda \in \rho(\mathcal{A})$. That is to say the spectrum of $\mathcal{A}$ is a pure point spectrum and $\sigma(\mathcal{A}) = \sigma(\Phi)$.

Next, given $\forall \lambda_0 \in \rho(\mathcal{A})$, we calculate $(\lambda_0 \mathcal{I} - \mathcal{A})^{-1} \mathcal{B}$. $\forall u \in \mathcal{U}$,

$$\mathcal{J}u \triangleq (\lambda_0 \mathcal{I} - \mathcal{A})^{-1} \mathcal{B}u = \left((\lambda_0 I - \Phi)^{-1} \sum_{i=0}^l e^{-\lambda_0 d_i} \Gamma_{(i)} u, \sum_{i=0}^l e^{-\lambda_0 d_i} \Gamma_{(i)} u \chi_{[-d_l, \tau)}(-d_i) \right).$$

From the above equation, we see that $\mathcal{J}$ is a linear bounded operator for all $\lambda_0 \in \rho(\mathcal{A})$. □

The cost function can be reformulated as

$$J_\infty = \lim_{T \rightarrow \infty} \int_0^T \left(\langle \phi(\tau), \mathcal{Q}(\phi(\tau)) \rangle_{\mathcal{M}_2} + u'(\tau) R u(\tau) \right) d\tau, \quad (4.47)$$

where $\mathcal{Q}$ is a continuous linear map on $\mathcal{M}_2$ and defined as

$$\mathcal{Q}(x, y) = (Qx, 0), \quad \forall (x, y) \in \mathcal{M}_2. \quad (4.48)$$

By reformulating the LQR problem, we have the following result.

Theorem 4.4. *For system (4.40) and cost function (4.47), there exists a family $\{\mathcal{P}(t), 0 \leq t \leq T\}$ of continuous linear map on $\mathcal{M}_2$ satisfying the following operator Riccati differential equation,*

$$-\dot{\mathcal{P}}(t) = \mathcal{A}^* \mathcal{P}(t) + \mathcal{P}(t) \mathcal{A} + \mathcal{Q} - \mathcal{K}(t) R \mathcal{K}^*(t), \quad 0 \leq t \leq T, \quad (4.49)$$

where $\mathcal{K}(t) = \mathcal{P}(t) \mathcal{B} R^{-1}$, $\mathcal{P}(T) = \mathcal{P}_T$ with $\mathcal{P}_T$ defined as $\mathcal{P}_T(x, y) = (P_T x, 0)$.

Furthermore, if $(\Phi, \sum_{i=0}^l e^{-\Phi d_i} \Gamma_{(i)})$ is stabilizable and $(Q^{\frac{1}{2}}, \Phi)$ is detectable, the Riccati differential equation (4.49) converges to $\mathcal{P}_\infty$ as $T \rightarrow \infty$, which is the solution to the following algebraic Riccati equation

$$\mathcal{A}^* \mathcal{P} + \mathcal{P} \mathcal{A} + \mathcal{Q} - \mathcal{P} \mathcal{B} R^{-1} \mathcal{B}^* \mathcal{P} = 0, \quad (4.50)$$

and the closed loop operator $\mathcal{F} = \mathcal{A} - \mathcal{B} R^{-1} \mathcal{B}^* \mathcal{P}_\infty$ is exponentially stable.

Proof. The proof of the first part uses standard arguments and will be omitted. What we need to prove is the convergence of (4.49) and the stability of the closed system under the stabilizable and detectable assumption.

If we write $\mathcal{Q}$ in the form of $\mathcal{C}^* \mathcal{C}$ where $\mathcal{C}$ is defined as $\mathcal{C}(x, y) = (Q^{\frac{1}{2}} x, 0)$, then we will show the stabilizability of $(\mathcal{A}, \mathcal{B})$ and detectability of $(\mathcal{C}, \mathcal{A})$. According to [8], we can prove it by contradiction. Assume that there exist a complex number η with

$Re\{\eta\} \geq 0$ and $(z, w) \neq 0 \in \mathcal{D}(\mathcal{A})$ satisfying

$$(\eta\mathcal{I} - \mathcal{A})(z, w) = 0, \quad \mathcal{C}(z, w) = 0. \quad (4.51)$$

In light of (4.45)-(4.46), we can obtain $\Phi z = \eta z$ and $Q^{\frac{1}{2}}z = 0$, which contradicts with the assumption of detectability of $(Q^{\frac{1}{2}}, \Phi)$. So we have proved that $(\mathcal{C}, \mathcal{A})$ is detectable. The stabilizability of $(\mathcal{A}, \mathcal{B})$ is equivalent to $\forall \lambda \in \sigma(\mathcal{A}), Re\{\lambda\} \geq 0, Ker(\lambda\mathcal{I} - \mathcal{A}^*) \cap Ker(\mathcal{B}^*) = \{0\}$, where $Ker(\cdot)$ is the null space. It is easy to check that if $(\Phi, \sum_{i=0}^l e^{-\Phi d_i} \Gamma_{(i)})$ is stabilizable, there will not exist η with $Re\{\eta\} \geq 0$ and $(z, w) \neq 0 \in \mathcal{D}(\mathcal{A}^*)$ satisfying

$$(\eta\mathcal{I} - \mathcal{A}^*)(z, w) = 0, \quad \mathcal{B}^*(z, w) = 0, \quad (4.52)$$

which means $(\mathcal{A}, \mathcal{B})$ is stabilizable. Since the spectrum of $\mathcal{A}$ is point spectrum and $\mathcal{B}$ can be expressed in the form of (4.44), it follows that the Riccati differential equation (4.49) converges by recalling [8], meanwhile, the operator $\mathcal{F}$ is exponentially stable. Furthermore, the closed-loop system (4.40), by setting $u(s) = -R^{-1}\mathcal{B}^*\mathcal{P}_\infty\phi(s)$, is asymptotically stable as $T \rightarrow \infty$. $\square$

By some tedious but straightforward calculations, the matrix function $P(t - \theta_1, t - \theta_2, t)$ that appears in (4.24)-(4.27) is the kernel of $\mathcal{P}(t)$. According to [23], we can easily prove that $P(t - \theta_1, t - \theta_2, t)$ is bounded with respect to (t, θ_1, θ_2) over $[0, \infty) \times [-d_l, 0] \times [-d_l, 0]$. On the other hand, the stability of the closed-loop system (4.40) implies the stability of the closed-loop system (4.38). Therefore, under the standard assumption of stabilizability and detectability, we know that the finite horizon LQR solution can be extended to the infinite horizon case. The steady-state solutions of the PDEs (4.24)-(4.27) are simplified as

$$\lim_{t \rightarrow \infty} P(t - \theta_1, t - \theta_2, t) = P(-\theta_1, -\theta_2), \quad -d_l \leq \theta_1 \leq 0, \quad -d_l \leq \theta_2 \leq 0, \quad (4.53)$$

and the corresponding

$$\lim_{t \rightarrow \infty} K(t - \theta, t) = K(-\theta), \quad -d_l \leq \theta \leq 0.$$

Then we have the following result.

Theorem 4.5. *Consider system (4.38) and cost function (4.39). If $(\Phi, \sum_{i=0}^l e^{-\Phi d_i} \Gamma_{(i)})$ is stabilizable and $(Q^{\frac{1}{2}}, \Phi)$ is detectable, the LQR controller $u^*(s)$ exists and is given by*

$$u^*(s) = -K'(0)x(s) - \int_{-d_l}^0 K'(-\tau) \tilde{u}_s^*(s - \tau) d\tau,$$

where $K(-\theta), -d_l \leq \theta \leq 0$ can be calculated by $K(-\theta) = \sum_{i=0}^l P(-\theta, d_i) \Gamma_{(i)} R^{-1}$ and $P(-\theta_1, -\theta_2)$ satisfies,

$$\Phi' P(0, 0) + P(0, 0) \Phi + Q = K(0) R K'(0), \quad (4.54)$$

$$\frac{\partial P(-\theta, 0)}{\partial \theta} = -P(-\theta, 0) \Phi + K(-\theta) R K'(0), \quad (4.55)$$

$$\left[\frac{\partial}{\partial \theta_1} + \frac{\partial}{\partial \theta_2} \right] P(-\theta_1, -\theta_2) = K(-\theta_1) R K'(-\theta_2). \quad (4.56)$$

The minimum cost function is given by

$$J_\infty^* = x_0' P(0, 0) x_0 + 2 \int_0^{d_l} x_0' P(0, \theta) \tilde{u}_0(\theta) d\theta + \int_0^{d_l} \int_0^{d_l} \tilde{u}_0'(\theta_1) P(\theta_1, \theta_2) \tilde{u}_0(\theta_2) d\theta_1 d\theta_2.$$

Proof. The proof is straightforward by extending the finite horizon results and considering Theorem 4.4. $\square$

Remark 4.8. *We have proved that the solutions of PDEs (4.24)-(4.27) will converge to the solutions of (4.54)-(4.56). Therefore, one way of finding $P(-\theta_1, -\theta_2)$ is to solve the steady state solution of (4.24)-(4.27) according to Remark 4.4.*

4.3 Conclusion

We have investigated the LQR problem for continuous-time systems with multiple delays in a single input channel. An explicit optimal controller is given in terms of the solution of a partial Riccati type differential equation for continuous-time systems. The stability analysis has also been provided for the infinite horizon case. Our approach can be extended to solve other control problems such as H_∞ control for systems with multiple delays in a single input channel. Further investigation can also be extended for systems with random input delay of known distribution.

Chapter 5

Linear Quadratic Tracking for Discrete-time Systems with Multiple Delays in a Single Input Channel

In this chapter, we study the optimal tracking problem for linear discrete-time systems with known multiple delays in single input channel which can be encountered in many applications such as traffic control, irrigation systems, economic systems, and networked control systems [105], [52], [139]. The class of systems under investigation contains those in [92], [71] and [134] as special cases. The tracking problem will be structured such that the state vector $x(k)$ will be driven as close as possible to a desired state vector $\bar{x}(k)$ under the requirement that the control input $u(k)$ does not deviate too much from a given control input $\bar{u}(k)$.

Using a reorganized output approach, we first decouple the input signals at various delayed time instants and establish a duality between the linear quadratic tracking problem and a state and white noise smoothing problem for an associated backward stochastic system. The optimal estimation problem can be solved via projection theory. The optimal tracking control is then given in terms of the current state,

previous inputs, and the output of an auxiliary backward deterministic delay system which is formulated to simplify the solution to the optimal tracking control.

5.1 Problem Statement

We consider the discrete-time linear system with multiple delays in a single input channel which is studied in Chapter 3 (3.1). For convenient reference, we rewrite it as follows

$$x(k+1) = \Phi(k)x(k) + \sum_{i=0}^l \Gamma_{(i)}(k)u(k-d_i), \quad (5.1)$$

where $x \in \mathbb{R}^n$ is the state with initial value $x(0) = x_0$, $u \in \mathbb{R}^m$ is the control input with initial values $u(k) = u_0(k)$ when $k < 0$, $\Phi(k)$ and $\Gamma_{(i)}(k)$, $i = 0, 1, \dots, l$ are bounded time-varying matrices with appropriate dimension. The delays are assumed to be exactly known nonnegative integers and, without loss of generality, of an increasing order: $0 = d_0 < d_1 < \dots < d_l = d$.

Consider the following quadratic performance index for the system (5.1):

$$\begin{aligned} J_N = & [x_{N+1} - \bar{x}_{N+1}]' P_{N+1} [x_{N+1} - \bar{x}_{N+1}] + \sum_{i=0}^N [u(i) - \bar{u}(i)]' R_i [u(i) - \bar{u}(i)] \\ & + \sum_{i=0}^N [x(i) - \bar{x}(i)]' Q_i [x(i) - \bar{x}(i)], \end{aligned} \quad (5.2)$$

where $N > d$ is an integer, x_{N+1} is the terminal state, i.e., $x_{N+1} = x(N+1)$; $\bar{x}_{N+1}$ and $\bar{x}(\cdot)$ are the state vector to be tracked, $\bar{u}(\cdot)$ are the ‘nominal’ values of the control input signal; $P_{N+1} \geq 0$ is the penalty weighting matrix on the terminal state, $R_i > 0$, $Q_i \geq 0$, $i = 0, \dots, N$ are weighting matrices on the variations of the input signal $u(i)$ from the nominal values and the state tracking error $x(i) - \bar{x}(i)$, respectively.

The linear quadratic tracking problem is to find the state feedback control $\{u^*(s), 0 \leq s \leq N\}$ such that the cost function J_N is minimized.

Remark 5.1. *Note that the MPC provides a good alternative for solving the linear quadratic tracking problem [62], especially for constrained optimal control problems. However, for the delay system (5.1), to apply the MPC technique, state augmentation is unavoidable and in many cases, MPC does not provide a recursive algorithm for the control law which requires a much higher computational cost as compared to the Riccati recursion based solution to be developed in this chapter. Furthermore, our approach will provide useful insights to the optimal control problems for systems with input delays.*

5.2 Main Result

We shall present our main result on the linear quadratic tracking problem. The derivation of the result will be presented in the next section.

First, associated with the system (5.1) and the cost function (5.2), we introduce the following deterministic backward delay system which is needed for computing the optimal tracking control law:

$$r(k) = \Phi'(k)r(k+1) - K_0(k)z(k) + Q(k)\bar{x}(k), \quad (5.3)$$

$$\begin{aligned} z(k) = & \sum_{i=0}^l \Gamma'_{(i)}(k+d_i)r(k+1+d_i)\chi_{[k,N]}(k+d_i) + R(k)\bar{u}(k) \\ & - \sum_{i=1}^l \sum_{j=1}^{d_i} \Gamma'_{(i)}(k+d_i)K_{d_i-j+1}(k+j)z(k+j)\chi_{[k,N]}(k+d_i), \end{aligned} \quad (5.4)$$

where the initial state $r(N+1) = P_{N+1}\bar{x}_{N+1}$. By defining $h_k = (d-1) \wedge (N-k)$, $K_i(k)$, $i = 0, \dots, h_k + 1$ in system (5.3)-(5.4) can be calculated step by step as follows:

- 1). Set $k = N$, $P(k+1, k+1, k+1) = P_{N+1}$.
- 2). Calculate $R_e(k)$ by

$$R_e(k) = \sum_{i,j=0}^l [\Gamma'_{(i)}(k+d_i)P(k+d_i+1, k+d_j+1, k+1)\Gamma_{(j)}(k+d_j)$$

$$\times \chi_{[k,N]}(k + d_i) \chi_{[k,N]}(k + d_j)] + R(k), \quad (5.5)$$

3). Calculate $K_i(k)$ using

$$K_i(k) = \sum_{j=0}^l [P(k + i, k + d_j + 1, k + 1) \Gamma_{(j)}(k + d_j) \chi_{[k,N]}(k + d_j)] R_e^{-1}(k),$$

$$i = 1, \dots, h_k + 1, \quad (5.6)$$

$$K_0(k) = \Phi'(k) K_1(k). \quad (5.7)$$

4). Update $P(k + i, k + j, k)$, $i, j = 0, \dots, h_k + 1$ in the following way

$$P(k + i, k + j, k) = P(k + i, k + j, k + 1) - K_i(k) R_e(k) K_j'(k),$$

$$i = 1, \dots, j, \quad j = 1, \dots, h_k + 1, \quad (5.8)$$

$$P(k, k + j, k) = \Phi'(k) P(k + 1, k + j, k), \quad j = 1, \dots, h_k + 1, \quad (5.9)$$

$$P(k, k, k) = \Phi'(k) P(k + 1, k + 1, k) \Phi(k) + Q(k), \quad (5.10)$$

$$P(k + i, k + j, k) = P'(k + j, k + i, k). \quad (5.11)$$

5). $k - 1 \Rightarrow k$. If $k \geq 0$, go back to 2); else, end.

Denote

$$\tilde{u}_s(k) \triangleq \sum_{j=1}^l \Gamma_{(j)}(k) u(k - d_j) \chi_{[-d,s]}(k - d_j), \quad k \geq s \geq 0, \quad (5.12)$$

which is a combination of the past control inputs. The main result is given below whose proof is deferred to the next section.

Theorem 5.1. *For system (5.1) and quadratic cost function (5.2), the optimal tracking controller $\{u^*(s), 0 \leq s \leq N\}$ that minimizes (5.2) can be calculated by*

$$u^*(s) = R_e^{-1}(s) z(s) - K_0'(s) x(s) - \sum_{i=0}^{h_s} K_{i+1}'(s) \tilde{u}_s^*(s + i), \quad (5.13)$$

where $z(s)$, $K_i(s)$, $i = 0, \dots, h_s + 1$ and $R_e(s)$ are as defined in the auxiliary backward

system (5.3)-(5.4), (5.5) and (5.6), respectively, $\tilde{u}_s^*(k)$ is given in (5.12) with $u_s(k - d_j)$ replaced by $u_s^*(k - d_j)$.

Remark 5.2. *Theorem 5.1 presents an analytical solution to the linear quadratic tracking problem for systems with multiple delays in single input channel. The solution is given in terms of the current state $x(s)$, the previous control inputs $\tilde{u}_s^*(s + i)$, $i = 0, 1, \dots, h_s$, and the output $z(s)$ of the auxiliary backward system (5.3)-(5.4). The feedback gain matrices $K_i(s)$, $i = 0, 1, \dots, h_s + 1$ and $R_e(s)$ rely on a Riccati difference recursion of the same dimension as the plant (ignoring the delays). It can be verified in Section 5.4 that the proposed solution is computationally much more efficient than that given by the state augmentation approach.*

5.3 Proof of Main Result

We shall prove the main result by introducing a dual backward stochastic system and establishing a duality between the optimal quadratic tracking problem and a smoothing problem for the backward stochastic system. Some elementary tools such as projection theory can be used for solving the smoothing problem for the backward stochastic system.

In order to derive a closed-form solution, we consider the delay system (5.1) starting from an arbitrary given time $0 \leq s \leq N$ rather than 0. By introducing the same notation as (3.3)-(3.7), $\forall k \geq s$, system (5.1) can be rewritten as

$$x(k+1) = \Phi(k)x(k) + \Gamma_s(k)u_s(k) + \tilde{u}_s(k). \quad (5.14)$$

The cost function (5.2) can be split into two terms,

$$J_N = J_N^s + \bar{J}_N^s,$$

where

$$\begin{aligned}
J_N^s &= [x_{N+1} - \bar{x}_{N+1}]' P_{N+1} [x_{N+1} - \bar{x}_{N+1}] + \sum_{i=s}^N [u(i) - \bar{u}(i)]' R(i) [u(i) - \bar{u}(i)] \\
&\quad + \sum_{i=s}^N [x(i) - \bar{x}(i)]' Q(i) [x(i) - \bar{x}(i)] \\
&= [x_{N+1} - \bar{x}_{N+1}]' P_{N+1} [x_{N+1} - \bar{x}_{N+1}] + \sum_{i=s}^N [u_s(i) - \bar{u}_s(i)]' R_s(i) [u_s(i) - \bar{u}_s(i)] \\
&\quad + \sum_{i=s}^N [x(i) - \bar{x}(i)]' Q(i) [x(i) - \bar{x}(i)], \tag{5.15}
\end{aligned}$$

$$\bar{J}_N^s = \sum_{i=0}^{s-1} [u(i) - \bar{u}(i)]' R(i) [u(i) - \bar{u}(i)] + \sum_{i=0}^{s-1} [x(i) - \bar{x}(i)]' Q(i) [x(i) - \bar{x}(i)]. \tag{5.16}$$

Note that $\{u(k), k \geq s\}$ does not affect $\bar{J}_N^s$. So, we shall design $\{u(k), k \geq s\}$ to minimize J_N^s .

Remark 5.3. *We note that although system (5.14) and cost function J_N^s in (5.15) appear to be of the form of a standard quadratic tracking problem, direct application of standard quadratic tracking result is not workable in view of the fact that $u_s(k)$ of (3.3) involves multiple delays and more importantly there exists a constraint that each element of $u_s(k)$ is from the same control input with different delays.*

5.3.1 Duality between Tracking and Smoothing Problems

Motivated by [135], we will convert the optimal tracking problem into a state and white noise smoothing problem for an associated backward stochastic state-space system. In view of the system model (5.14) and the cost function J_N^s of (5.15), we introduce the following associated backward stochastic system:

$$\mathbf{x}(k) = \Phi'(k) \mathbf{x}(k+1) + \mathbf{q}(k), \tag{5.17}$$

$$\mathbf{y}(k) = \Gamma'_s(k) \mathbf{x}(k+1) + \mathbf{v}(k), \quad s \leq k \leq N, \tag{5.18}$$

where $\mathbf{x}(N+1) = \mathbf{x}_{N+1}$, $\mathbf{q}(k)$ and $\mathbf{v}(k)$ are uncorrelated white noises with zero means and covariances $\langle \mathbf{x}_{N+1}, \mathbf{x}_{N+1} \rangle = P_{N+1}$, $\langle \mathbf{q}(i), \mathbf{q}(j) \rangle = Q_i \delta_{i,j}$, $\langle \mathbf{v}(i), \mathbf{v}(j) \rangle = R_s(i) \delta_{i,j}$.

Remark 5.4. *In the following, we shall show that the linear quadratic tracking problem for system (5.14) is dual to the state and white noise smoothing problem for backward stochastic system (5.17)-(5.18). However, we cannot directly solve the tracking problem based on this duality since it will lead to a constrained optimization problem (due to the fact that $u_s(k)$ contains elements of the same input with different delays). We will derive a new backward stochastic system corresponding to system (5.1) and cost function (5.15) based on system (5.17)-(5.18).*

After introducing system (5.17)-(5.18), we will first rewrite the cost function J_N^s of (5.15) in terms of the parameters of the backward stochastic system (5.17)-(5.18).

Lemma 5.1. *By making use of the stochastic state-space model (5.17)-(5.18), J_N^s of (5.15) can be put in the following quadratic form*

$$J_N^s = \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}' \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix} \right\rangle \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}, \quad (5.19)$$

where $\xi_s = [x'(s) \ \tilde{u}'_s(s) \ \cdots \ \tilde{u}'_s(s+h_s)]'$, with $\tilde{u}_s(s+i)$, $i = 0, \dots, h_s$ as defined in (5.12), and

$$\begin{aligned} \bar{x}_s &= \text{col}\{\bar{x}(s), \dots, \bar{x}(N)\}, \quad \bar{u}_s = \text{col}\{\bar{u}_s(s), \dots, \bar{u}_s(N)\}, \quad u_s = \text{col}\{u_s(s), \dots, u_s(N)\}, \\ \mathbf{q}_s &= \text{col}\{\mathbf{q}(s), \dots, \mathbf{q}(N)\}, \quad \mathbf{v}_s = \text{col}\{\mathbf{v}(s), \dots, \mathbf{v}(N)\}, \quad \mathbf{y}_s = \text{col}\{\mathbf{y}(s), \dots, \mathbf{y}(N)\}, \\ \mathbf{x}_s^0 &= \text{col}\{\mathbf{x}(s), \mathbf{x}(s+1), \dots, \mathbf{x}(s+h_s+1)\}. \end{aligned}$$

Proof. According to system (5.14), we have

$$x_s = \mathcal{O}_{N,s}\xi_s + \mathcal{B}_{N,s}u_s, \quad x_{N+1} = \mathcal{H}_{N,s}\xi_s + \mathcal{C}_{N,s}u_s, \quad (5.20)$$

where

$$x_s = \text{col} \{x(s), \dots, x(N)\}, \quad (5.21)$$

$$\mathcal{O}_{N,s} = \begin{bmatrix} I_n & 0 & \cdots & 0 \\ \Pi_{s,s}^\Phi & I_n & \ddots & \vdots \\ \Pi_{s+1,s}^\Phi & \Pi_{s+1,s+1}^\Phi & \ddots & 0 \\ \vdots & \vdots & \ddots & I_n \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{N-1,s}^\Phi & \Pi_{N-1,s+1}^\Phi & \cdots & \Pi_{N-1,s+d}^\Phi \end{bmatrix}, \quad (5.22)$$

$$\mathcal{B}_{N,s} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \Gamma_s(s) & 0 & \cdots & 0 \\ \Pi_{s+1,s+1}^\Phi \Gamma_s(s) & \Gamma_s(s+1) & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{N-1,s+1}^\Phi \Gamma_s(s) & \Pi_{N-1,s+2}^\Phi \Gamma_s(s+1) & \cdots & \Gamma_s(N-1) \end{bmatrix}, \quad (5.23)$$

$$\mathcal{H}_{N,s} = [\Pi_{N,s}^\Phi \quad \Pi_{N,s+1}^\Phi \quad \cdots \quad \Pi_{N,s+d}^\Phi], \quad (5.24)$$

$$\mathcal{C}_{N,s} = [\Pi_{N,s+1}^\Phi \Gamma_s(s) \quad \Pi_{N,s+2}^\Phi \Gamma_s(s+1) \quad \cdots \quad \Pi_{N,N}^\Phi \Gamma_s(N-1) \quad \Gamma_s(N)]. \quad (5.25)$$

Then by substituting x_s and x_{N+1} into J_N^s , we can write J_N^s in the following quadratic form

$$\begin{aligned} J_N^s &= (\mathcal{H}_{N,s}\xi_s + \mathcal{C}_{N,s}u_s - \bar{x}_{N+1})' P_{N+1} (\mathcal{H}_{N,s}\xi_s + \mathcal{C}_{N,s}u_s - \bar{x}_{N+1}) \\ &\quad + (u_s - \bar{u}_s)' \bar{R}_s (u_s - \bar{u}_s) + (\mathcal{O}_{N,s}\xi_s + \mathcal{B}_{N,s}u_s - \bar{x}_s)' \bar{Q}_s (\mathcal{O}_{N,s}\xi_s + \mathcal{B}_{N,s}u_s - \bar{x}_s) \end{aligned}$$

$$= \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}' \begin{bmatrix} \Omega_1 & * & * & * & * \\ -P_{N+1}\mathcal{H}_{N,s} & P_{N+1} & * & * & * \\ -\bar{Q}_s\mathcal{O}_{N,s} & 0 & \bar{Q}_s & * & * \\ 0 & 0 & 0 & \bar{R}_s & * \\ \Omega_2 & -\mathcal{C}'_{N,s}P_{N+1} & -\mathcal{B}'_{N,s}\bar{Q}_s & -\bar{R}_s & \Omega_3 \end{bmatrix} \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}, \quad (5.26)$$

where

$$\begin{aligned} \bar{Q}_s &= \text{diag}\{Q_s, \dots, Q_N\}, \quad \bar{R}_s = \text{diag}\{R_s(s), \dots, R_s(N)\}, \\ \Omega_1 &= \mathcal{H}'_{N,s}P_{N+1}\mathcal{H}_{N,s} + \mathcal{O}'_{N,s}\bar{Q}_s\mathcal{O}_{N,s}, \quad \Omega_2 = \mathcal{C}'_{N,s}P_{N+1}\mathcal{H}_{N,s} + \mathcal{B}'_{N,s}\bar{Q}_s\mathcal{O}_{N,s}, \\ \Omega_3 &= \mathcal{C}'_{N,s}P_{N+1}\mathcal{C}_{N,s} + \mathcal{B}'_{N,s}\bar{Q}_s\mathcal{B}_{N,s} + \bar{R}_s. \end{aligned}$$

On the other hand, in light of (5.17)-(5.18), we write $\mathbf{x}_s^0$ and $\mathbf{y}_s$ in the following compact form

$$\mathbf{x}_s^0 = \mathcal{H}'_{N,s}\mathbf{x}_{N+1} + \mathcal{O}'_{N,s}\mathbf{q}_s, \quad \mathbf{y}_s = \mathcal{C}'_{N,s}\mathbf{x}_{N+1} + \mathcal{B}'_{N,s}\mathbf{q}_s + \mathbf{v}_s. \quad (5.27)$$

It is easy to verify that the center matrix of (5.26), denoted by Π_s , can be expressed in the following form

$$\Pi_s = \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix} \right\rangle,$$

which completes the proof. $\square$

Since the elements of $u_s(k)$ are obtained from the same input with different delays, we apply a re-organization of measured outputs to overcome the difficulty arising

from this constraint. First, decompose $\mathbf{y}(k)$ and $\mathbf{v}(k)$ of (5.18) as follows:

$$\mathbf{y}(k) = \begin{cases} \text{col} \{ \mathbf{y}_0(k), \dots, \mathbf{y}_i(k) \}, & d_i \leq k - s < d_{i+1}, \\ \text{col} \{ \mathbf{y}_0(k), \dots, \mathbf{y}_l(k) \}, & k - s \geq d, \end{cases} \quad (5.28)$$

$$\mathbf{v}(k) = \begin{cases} \text{col} \{ \mathbf{v}_0(k), \dots, \mathbf{v}_i(k) \}, & d_i \leq k - s < d_{i+1}, \\ \text{col} \{ \mathbf{v}_0(k), \dots, \mathbf{v}_l(k) \}, & k - s \geq d, \end{cases} \quad (5.29)$$

where $\mathbf{y}_i(k)$ and $\mathbf{v}_i(k)$, $i = 0, \dots, l$ satisfy

$$\mathbf{y}_i(k) = \Gamma'_{(i)}(k) \mathbf{x}(k+1) + \mathbf{v}_i(k), \quad (5.30)$$

with $\langle \mathbf{v}_i(k_1), \mathbf{v}_j(k_2) \rangle = R_{i,k_1} \delta_{i,j} \delta_{k_1,k_2}$. By using the same manipulation as that in Chapter 3, we introduce the following output

$$\bar{\mathbf{y}}(k) = \sum_{j=0}^l \mathbf{y}_j(k + d_j) \chi_{[k,N]}(k + d_j). \quad (5.31)$$

In view of (5.30), $\bar{\mathbf{y}}(k)$ can be rewritten as

$$\bar{\mathbf{y}}(k) = \sum_{j=0}^l \Gamma'_{(j)}(k + d_j) \mathbf{x}(k + d_j + 1) \chi_{[k,N]}(k + d_j) + \bar{\mathbf{v}}(k), \quad (5.32)$$

where

$$\bar{\mathbf{v}}(k) = \sum_{j=0}^l \mathbf{v}_j(k + d_j) \chi_{[k,N]}(k + d_j). \quad (5.33)$$

It is easy to verify that $\bar{\mathbf{v}}(k)$, $k = s, \dots, N$ are uncorrelated white noises with zero means and covariances $\langle \bar{\mathbf{v}}(i), \bar{\mathbf{v}}(k) \rangle = R_i \delta_{i,k}$. Then, we arrive at the following result.

Lemma 5.2. *Let*

$$\begin{aligned} U_s &= \text{col} \{ u(s), \dots, u(N) \}, \quad \bar{U}_s = \text{col} \{ \bar{u}(s), \dots, \bar{u}(N) \}, \\ \bar{\mathbf{y}}_s &= \text{col} \{ \bar{\mathbf{y}}(s), \dots, \bar{\mathbf{y}}(N) \}, \quad \bar{\mathbf{v}}_s = \text{col} \{ \bar{\mathbf{v}}(s), \dots, \bar{\mathbf{v}}(N) \}, \end{aligned}$$

where $\bar{\mathbf{y}}(k)$ is as in (5.32) and $\bar{\mathbf{v}}(k)$ in (5.33). Then J_N^s of (5.15) can be further

rewritten in the following quadratic form

$$J_N^s = \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{U}_s \\ U_s \end{bmatrix}' \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\bar{\mathbf{v}}_s \\ \bar{\mathbf{y}}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_{N+1} \\ -\mathbf{q}_s \\ -\bar{\mathbf{v}}_s \\ \bar{\mathbf{y}}_s \end{bmatrix} \right\rangle \begin{bmatrix} \xi_s \\ \bar{x}_{N+1} \\ \bar{x}_s \\ \bar{U}_s \\ U_s \end{bmatrix}. \quad (5.34)$$

Furthermore, the optimal control is given by

$$U_s^* = R_{\bar{\mathbf{y}}_s}^{-1} (R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}} \bar{x}_{N+1} + R_{\bar{\mathbf{y}}_s \mathbf{q}_s} \bar{x}_s + R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}_s} \bar{U}_s - R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} \xi_s), \quad (5.35)$$

where $U_s^* = \text{col}\{u^*(s), \dots, u^*(N)\}$, $R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}} = \langle \bar{\mathbf{y}}_s, \mathbf{x}_{N+1} \rangle$, $R_{\bar{\mathbf{y}}_s \mathbf{q}_s} = \langle \bar{\mathbf{y}}_s, \mathbf{q}_s \rangle$, $R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} = \langle \bar{\mathbf{y}}_s, \mathbf{x}_s^0 \rangle$, $R_{\bar{\mathbf{y}}_s} = \langle \bar{\mathbf{y}}_s, \bar{\mathbf{y}}_s \rangle$.

Proof. The proof is omitted and a similar proof can be found in [136]. $\square$

Remark 5.5. Lemma 5.2 provides an explicit form of the optimal input U_s^* in terms of ξ_s , $\bar{x}_{N+1}$, $\bar{x}_s$ and $\bar{U}_s$. All of the gain matrices ($R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}}$, $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}_s}$, $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}_s}$ and $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0}$) can be obtained by solving the corresponding optimal estimation problem. In fact, the gain matrices for the estimation of $\mathbf{x}_{N+1}$, $\mathbf{q}_s$, $\bar{\mathbf{v}}_s$ and $\mathbf{x}_s^0$ based on the outputs $\bar{\mathbf{y}}_s$ are $R_{\mathbf{x}_{N+1} \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$, $R_{\mathbf{q}_s \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$, $R_{\bar{\mathbf{v}}_s \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$ and $R_{\mathbf{x}_s^0 \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$, respectively, are the transpose of the optimal controller gain matrices. Consequently, we solve the corresponding optimal estimation problem instead.

5.3.2 Formation of the Backward System and Solution to the Tracking Problem

Lemma 5.2 reveals a duality between the linear quadratic tracking problem and the optimal state and white noise smoothing estimation problem for the stochastic backward system (5.17) and (5.32). Theorem 3.1 provides a recursive calculation of the optimal estimate of $\mathbf{x}_s^0$, based on which we will derive the evolution equation

of the innovation $\mathbf{e}(k)$ in the next theorem which will be useful in establishing the result of Theorem 5.1.

Theorem 5.2. *For system (5.17) and (5.32), $\mathbf{e}(k)$ defined in Theorem 3.1, satisfies*

$$\begin{aligned} \mathbf{e}(k) = & \bar{\mathbf{v}}(k) + \sum_{i=0}^l \Gamma'_{(i)}(k + d_i) \tilde{\mathbf{x}}(k + d_i + 1 | k + d_i + 1) \chi_{[k, N]}(k + d_i) \\ & - \sum_{i=1}^l \sum_{j=1}^{d_i} \Gamma'_{(i)}(k + d_i) K_{d_i-j+1}(k + j) \mathbf{e}(k + j) \chi_{[k, N]}(k + d_i), \end{aligned} \quad (5.36)$$

$$\tilde{\mathbf{x}}(k | k) = \Phi'(k) \tilde{\mathbf{x}}(k + 1 | k + 1) - K_0(k) \mathbf{e}(k) + \mathbf{q}(k), \quad (5.37)$$

with the initial condition $\tilde{\mathbf{x}}(N + 1 | N + 1) = \mathbf{x}_{N+1}$.

Proof. It can be proved by combining (3.34), (3.35) and (3.46). $\square$

We have given a recursive method for calculating the optimal estimator of $\mathbf{x}_s^0$ in Theorem 3.1. Unfortunately, we cannot get an explicit form of the estimators of $\mathbf{x}_{N+1}$, $\mathbf{q}(k)$ and $\bar{\mathbf{v}}(k)$ by the projection approach due to the existence of multiple state delays in $\bar{\mathbf{y}}(k)$ of (5.32). Some further manipulations need to be done. In the following, we will formulate an auxiliary backward deterministic system to overcome this difficulty.

Because the system (5.36)-(5.37) is linear, we shall rewrite $\mathbf{e}(s)$ in the following way

$$\mathbf{e}(s) = \Psi_{N+1,s}^x \mathbf{x}_{N+1} + \sum_{k=s+1}^N \Psi_{k,s}^q \mathbf{q}(k) + \sum_{k=s}^N \Psi_{k,s}^v \bar{\mathbf{v}}(k), \quad (5.38)$$

where $\Psi_{N+1,s}^x$, $\Psi_{k,s}^q$ and $\Psi_{k,s}^v$ are the transition matrices from $\mathbf{x}_{N+1}$, $\mathbf{q}(k)$ and $\bar{\mathbf{v}}(k)$ to $\mathbf{e}(s)$, respectively. Due to the independence of $\mathbf{x}_{N+1}$, $\mathbf{q}(k)$ and $\bar{\mathbf{v}}(k)$, we have

$$R_{\mathbf{e}(s), \mathbf{x}_{N+1}} \triangleq \langle \mathbf{e}(s), \mathbf{x}_{N+1} \rangle = \Psi_{N+1,s}^x P_{N+1}, \quad (5.39)$$

$$R_{\mathbf{e}(s), \mathbf{q}(k)} \triangleq \langle \mathbf{e}(s), \mathbf{q}(k) \rangle = \Psi_{k,s}^q Q_k, \quad k > s, \quad (5.40)$$

$$R_{\mathbf{e}(s), \bar{\mathbf{v}}(k)} \triangleq \langle \mathbf{e}(s), \bar{\mathbf{v}}(k) \rangle = \Psi_{k,s}^v R_k, \quad k \geq s. \quad (5.41)$$

Next, we will derive an explicit form of the gain matrix from $\bar{x}_{N+1}$, $\bar{x}(k)$ and $\bar{u}(k)$ to U_s^* . According to the projection theory,

$$\begin{aligned}\bar{\mathbf{y}}'_s R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}} &= \sum_{k=s}^N \mathbf{e}'(k) R_e^{-1}(k) R_{\mathbf{e}(s), \mathbf{x}_{N+1}} \\ &= (\bar{\mathbf{y}}(s) - \hat{\bar{\mathbf{y}}}(\mathbf{s}))' R_e^{-1}(s) R_{\mathbf{e}(s), \mathbf{x}_{N+1}} + \sum_{k=s+1}^N \mathbf{e}'(k) R_e^{-1}(k) R_{\mathbf{e}(s), \mathbf{x}_{N+1}} \\ &= \bar{\mathbf{y}}'(s) R_e^{-1}(s) \Psi_{N+1,s}^x P_{N+1} + \bar{\mathbf{y}}'_{s+1} \bar{K}_{N+1,s+1}^x,\end{aligned}\quad (5.42)$$

where $\hat{\bar{\mathbf{y}}}(\mathbf{s})$ is the projection of $\bar{\mathbf{y}}(s)$ onto the linear space $\mathcal{L}\{\bar{\mathbf{y}}_{s+1}\}$, and

$$\bar{\mathbf{y}}'_{s+1} \bar{K}_{N+1,s+1}^x \triangleq \sum_{k=s+1}^N \mathbf{e}'(k) R_e^{-1}(k) R_{\mathbf{e}(s), \mathbf{x}_{N+1}} - \hat{\bar{\mathbf{y}}}'(s) R_e^{-1}(s) \Psi_{N+1,s}^x P_{N+1}. \quad (5.43)$$

Hence

$$\begin{aligned}R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}} &= R_{\bar{\mathbf{y}}_s}^{-1} \langle \bar{\mathbf{y}}_s, \bar{\mathbf{y}}'_s R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_{N+1}} \rangle \\ &= R_{\bar{\mathbf{y}}_s}^{-1} \langle \bar{\mathbf{y}}_s, \bar{\mathbf{y}}'(s) R_e^{-1}(s) \Psi_{N+1,s}^x P_{N+1} + \bar{\mathbf{y}}'_{s+1} \bar{K}_{N+1,s+1}^x \rangle \\ &= \begin{bmatrix} R_e^{-1}(s) \Psi_{N+1,s}^x P_{N+1} \\ \bar{K}_{N+1,s+1}^x \end{bmatrix}.\end{aligned}\quad (5.44)$$

In fact, what we care is the first entry of the above matrix rather than the whole matrix since only the gain matrix from $\bar{x}_{N+1}$ to $u^*(s)$ rather than to U_s^* is useful to us. Analogously, by denoting

$$R_{\bar{\mathbf{y}}_s \mathbf{q}(k)} \triangleq \langle \bar{\mathbf{y}}_s, \mathbf{q}(k) \rangle, \quad R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(k)} \triangleq \langle \bar{\mathbf{y}}_s, \bar{\mathbf{v}}(k) \rangle,$$

we can write $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(k)}$ and $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(k)}$ in the following forms

$$R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(k)} = \begin{bmatrix} R_e^{-1}(s) \Psi_{k,s}^q Q_k \\ \bar{K}_{k,s+1}^q \end{bmatrix}, \quad (5.45)$$

$$R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(k)} = \begin{bmatrix} R_e^{-1}(s) \Psi_{k,s}^v R_k \\ \bar{K}_{k,s+1}^v \end{bmatrix}, \quad (5.46)$$

where $\bar{K}_{k,s+1}^q$ and $\bar{K}_{k,s+1}^v$ are defined in the same way as $\bar{K}_{N+1,s+1}^x$. Again using projection theory and considering Theorem 3.1, we have

$$\begin{aligned} \bar{\mathbf{y}}'_s R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} &= \sum_{k=s}^N \mathbf{e}'(k) R_e^{-1}(k) \langle \mathbf{e}(k), \mathbf{x}_s^0 \rangle \\ &= (\bar{\mathbf{y}}(s) - \hat{\mathbf{y}}(s))' R_e^{-1}(s) \langle \mathbf{e}(s), \mathbf{x}_s^0 \rangle + \sum_{k=s+1}^N \mathbf{e}'(k) R_e^{-1}(k) \langle \mathbf{e}(k), \mathbf{x}_s^0 \rangle \\ &= \bar{\mathbf{y}}'(s) K_s + \bar{\mathbf{y}}'_{s+1} \bar{K}_s, \end{aligned} \quad (5.47)$$

where $K_s = [K'_0(s) \ K'_1(s) \ \cdots \ K'_{h_s+1}(s)]$, and

$$\bar{\mathbf{y}}'_{s+1} \bar{K}_s \triangleq \sum_{k=s+1}^N \mathbf{e}'(k) R_e^{-1}(k) \langle \mathbf{e}(k), \mathbf{x}_s^0 \rangle - \hat{\mathbf{y}}'(s) K_s. \quad (5.48)$$

By applying similar manipulation, the gain matrix from ξ_s to U_s^* can be written as

$$R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{x}_s^0} = \begin{bmatrix} K_s \\ \bar{K}_s \end{bmatrix}.$$

Since

$$\begin{aligned} R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}_s} &= [R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(s)} \ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(s+1)} \ \cdots \ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(N)}], \ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \mathbf{q}(s)} = 0, \\ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}_s} &= [R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(s)} \ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(s+1)} \ \cdots \ R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s \bar{\mathbf{v}}(N)}], \end{aligned}$$

according to (5.35) in Lemma 5.2, the optimal input is given by

$$\begin{aligned} U_s^* &= - \begin{bmatrix} K_s \\ \bar{K}_s \end{bmatrix} \xi_s + \begin{bmatrix} R_e^{-1}(s) \Psi_{N+1,s}^x P_{N+1} \\ \bar{K}_{N+1,s+1}^x \end{bmatrix} \bar{x}_{N+1} \\ &\quad + \begin{bmatrix} R_e^{-1}(s) \Psi_{s+1,s}^q Q_{s+1} & R_e^{-1}(s) \Psi_{s+2,s}^q Q_{s+2} & \cdots & R_e^{-1}(s) \Psi_{N,s}^q Q_N \\ \bar{K}_{s+1,s+1}^q & \bar{K}_{s+2,s+1}^q & \cdots & \bar{K}_{N,s+1}^q \end{bmatrix} \bar{x}_{s+1} \end{aligned}$$

$$+ \begin{bmatrix} R_e^{-1}(s)\Psi_{s,s}^v R_s & R_e^{-1}(s)\Psi_{s+1,s}^v R_{s+1} & \cdots & R_e^{-1}(s)\Psi_{N,s}^v R_N \\ \bar{K}_{s+1,s+1}^v & \bar{K}_{s+2,s+1}^v & \cdots & \bar{K}_{N,s+1}^v \end{bmatrix} \bar{u}_s. \quad (5.49)$$

Note that the first row of (5.49) gives the optimal input at time instant s , i.e.

$$u^*(s) = -K_s \xi_s + R_e^{-1}(s) \left[\Psi_{N+1,s}^x P_{N+1} \bar{x}_{N+1} + \sum_{i=s+1}^N \Psi_{i,s}^q Q_i \bar{x}(i) + \sum_{j=s}^N \Psi_{j,s}^v R_j \bar{u}(j) \right]. \quad (5.50)$$

If we set the output of the system (5.36)-(5.37) as $R_e^{-1}(s)\mathbf{e}(s)$, we will realize that the second term on the right hand side of (5.50) is the output of system (5.36)-(5.37) by replacing $\mathbf{x}_{N+1}$, $\mathbf{q}(i)$ and $\bar{\mathbf{v}}(j)$ with $P_{N+1}\bar{x}_{N+1}$, $Q_i\bar{x}(i)$ and $R_j\bar{u}(j)$ for all $s < i \leq N$, $s \leq j \leq N$, respectively. Note that after the replacement, the system (5.36)-(5.37) becomes a deterministic system, which is (5.3)-(5.4).

Remark 5.6. *Note that (5.50) provides an explicit closed loop feedback form of the optimal controller. However, the transition matrices $\Psi_{N+1,s}^x$, $\Psi_{i,s}^q$ and $\Psi_{j,s}^v$ can not be calculated explicitly and therefore the controller can not be realized. We realized these transition matrices by introducing the auxiliary backward system (5.3)-(5.4) and all the controller gain matrices rely on the solution of the Riccati difference recursion (5.8)-(5.11).*

5.4 Analysis of Computational Complexity

Theorem 5.1 provides a solution to the optimal tracking problem for the delay system (5.1) via a non-augmentation approach. The computation of the optimal control law involves a Riccati difference recursion of the same dimension as the plant (ignoring the delays). The presented approach is of a significant computational advantage as compared with the augmentation method as demonstrated below.

Because the multiplications and divisions cost much more as compared to additions, we can use the number of multiplications and divisions as the operation count. Denote C_{aug} and C_{new} the numbers of the multiplications and divisions for the

augmentation method and the presented approach in one step, respectively. C_{aug} and C_{new} can be given as follows:

$$C_{aug} = 2m^3d^3 + (6m^2n + 4m^3 + 4m^2)d^2 + (6mn^2 + 8m^2n + 8mn + 3m^3 + 7m^2)d + 2n^3 + 4mn^2 + 5n^2 + 3m^2n + 7mn + m^3 + m^2, \quad (5.51)$$

$$C_{new} = 0.5(mn^2 + m^2n)d^2 + [n^3 + (l + 1.5)mn^2 + 1.5m^2n + mn]d + 3mn \sum_{i=1}^l d_i + n^3 + (l + 2)mn^2 + 2n^2 + (l + 1)m^2n + (l + 3)mn + m^3 + 2m^2. \quad (5.52)$$

From (5.51) and (5.52), it is clear that the order of d in C_{aug} is 3, while the order of d in C_{new} is 2. So when the delay is sufficiently large, it is easy to know that $C_{aug} \gg C_{new}$. To show this point, we consider the following example.

Example 5.1. Consider system (5.1) with $n = 3$, $m = 2$, and $l = 3$. Assume $d_i = i \cdot t$. As $i = 0, \dots, l$, $t = 1, \dots, 7$, the difference between C_{aug} and C_{new} is shown in Fig. 5.1.

Figure 5.1: Comparison of complexities between augmentation method and the presented approach

5.5 Applications

To illustrate the efficiency of the approach, we will present an application in economics, which is mainly motivated by [116].

Suppose that a firm produces two kinds of products. We need to determine how much should be spent on the advertisement of the products in order to increase the sales and how many products should be produced. The amount of sales of the two kinds of products in the k -th months are denoted by $S_1(k)$ and $S_2(k)$, respectively. For product $i = 1, 2$, we also denote by $A_i(k)$, $I_i(k)$ and $P_i(k)$ the advertisement cost, the amount of the inventory and the amount of the production, respectively. Let the sales of the product 1 (or 2) be affected by the last sales of the product 1 (or 2) as well as the previous advertisement. It is reasonable to assume that the advertisement of product 1 (or 2) has a delay effect on the sales of the product 2 (or 1). In that case, we can formulate it in the following way

$$\begin{cases} S_1(k+1) = 1.1S_1(k) + 5A_1(k) + 2A_2(k-2), \\ S_2(k+1) = 1.1S_2(k) + A_1(k-1) + 4A_2(k), \\ 0 \leq k \leq 11. \end{cases} \quad (5.53)$$

The relationship between inventory and production is given by

$$I_i(k+1) = I_i(k) + P_i(k+1) - S_i(k+1), \quad i = 1, 2. \quad (5.54)$$

All the variables here are assumed to be the incremental values from the desired values, and thus can be either positive or negative. Let

$$x(k) = [S_1(k) \ S_2(k) \ I_1(k) \ I_2(k)]', \quad u(k) = [\bar{P}_1(k) \ \bar{P}_2(k) \ A_1(k) \ A_2(k)]',$$

where $\bar{P}_1(k) = P_1(k+1)$, $\bar{P}_2(k) = P_2(k+1)$. The system can be described in the following compact form

$$x(k+1) = \Phi x(k) + \Gamma_{(0)}u(k) + \Gamma_{(1)}u(k-1) + \Gamma_{(2)}u(k-2), \quad (5.55)$$

where

$$\Phi = \begin{bmatrix} 1.1 & 0 & 0 & 0 \\ 0 & 1.1 & 0 & 0 \\ -1.1 & 0 & 1 & 0 \\ 0 & -1.1 & 0 & 1 \end{bmatrix}, \quad \Gamma_{(0)} = \begin{bmatrix} 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \\ 1 & 0 & -5 & 0 \\ 0 & 1 & 0 & -4 \end{bmatrix},$$

$$\Gamma_{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \Gamma_{(2)} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The initial condition is $x(0) = [5 \ 14 \ 10 \ 5]'$, $u(-2) = u(-1) = 0$. The cost function to be minimized is

$$J_{12} = [x(12) - \bar{x}(12)]' P_{12} [x(12) - \bar{x}(12)] + \sum_{i=0}^{11} [u(i) - \bar{u}(i)]' R [u(i) - \bar{u}(i)]$$

$$+ \sum_{i=0}^{11} [x(i) - \bar{x}(i)]' Q [x(i) - \bar{x}(i)],$$

where $R = \text{diag}\{5, 2.5, 2, 1\}$, $Q = \text{diag}\{1, 2, 1.5, 2.5\}$, $P_{12} = Q$.

By using Theorem 5.1, the optimal control strategies can be obtained. Fig. 5.2-Fig. 5.5 demonstrate that the variations of the sales, inventories, production outputs and advertisement costs from their desired values $\bar{x}(k) = [1.46 \ 1.265 \ 4.3218 \ 5.113]'$ and $\bar{u}(k) = [1.4589 \ 1.2655 \ -0.025 \ -0.028]'$ converge to zero fast.

5.6 Conclusion

We have investigated the linear quadratic tracking control problem for discrete-time systems with multiple delays in a single input channel. The optimal controller is given in terms of the solution of a Riccati difference equation and the output of an auxiliary backward delay system which is formulated for the first time. The proposed

Figure 5.2: Sales of products.

Figure 5.3: Inventory of products.

approach saves much computational cost comparing with traditional method. The numerical example illustrates the control design and demonstrates the effectiveness of the proposed approach.

Figure 5.4: Production of products.

Figure 5.5: Advertisement of products.

Chapter 6

Linear Quadratic Tracking for Continuous-time Systems with Multiple Delays in a Single Input Channel

6.1 Introduction

In this chapter, we shall study the linear quadratic tracking problem for linear continuous-time systems with multiple delays in a single input channel. Different from [71], the problem in this paper is very difficult to solve due to the coupling of the delayed input signals. The system to be considered in this chapter is widely encountered in practical systems. Take water resource distribution for an example. As is shown in Fig. 6.1, there are two reservoirs R_1 and R_2 . We have to control the diverted water u_2 to change the inflow into the irrigable land. A portion of diverted irrigation water αu_2 returns to the main stream with time delay. We can see that the change of u_1 will result in the change of the downstream inflow not only through the main stream but also through αu_2 with time delay.

Figure 6.1: Water distribution system

In fact, for systems considered in [71], each delayed input channel is independent of each other and their control law can be designed independently. However, the system to be considered in this chapter is that of [71] with the constraint that the inputs of all the channels must be the same. This constraint poses a significant challenge and cannot be addressed using the approach of [71] as it will lead to a constrained optimization problem to which no analytical solution can be obtained. By using a reorganized output approach, we establish a duality between the linear quadratic tracking problem and a state and white noise smoothing problem for an associated backward linear continuous stochastic system. The derived optimal estimation problem can be solved via the approach in Chapter 4.

6.2 Problem Statement

We consider the continuous-time system with multiple delays in a single input channel as is in Chapter 4. For easy reference, we rewrite it as follows

$$\dot{x}(t) = \Phi(t)x(t) + \sum_{i=0}^l \Gamma_{(i)}(t)u(t - d_i), \quad (6.1)$$

where $x \in \mathbb{R}^n$ is the state with initial status $x(0) = x_0$, $u \in \mathbb{R}^m$ is the input with initial values $u(t) = \mu_t$, $t < 0$. $\Phi(t)$ and $\Gamma_{(i)}(t)$, $i = 0, 1, \dots, l$ are bounded time-

varying matrices with appropriate dimension. The delays are assumed to be of an increasing order: $0 = d_0 < d_1 < \dots < d_l \triangleq d$.

Given a desired state trajectory $\bar{x}(t)$ and a ‘nominal’ value of input $\bar{u}(t)$, what we are trying to do is to minimize the ‘energy’ of the difference between the true value of $x(t)$, $u(t)$ and the desired value $\bar{x}(t)$, $\bar{u}(t)$. We consider the following quadratic cost function for system (6.1):

$$\begin{aligned} J_T = & (x_T - \bar{x}_T)' P_T (x_T - \bar{x}_T) + \int_0^T (x(t) - \bar{x}(t))' Q_t (x(t) - \bar{x}(t)) dt \\ & + \int_0^T (u(t) - \bar{u}(t))' R_t (u(t) - \bar{u}(t)) dt, \end{aligned} \quad (6.2)$$

where $x_T \triangleq x(T)$ and $\bar{x}_T \triangleq \bar{x}(T)$ are terminal values of the real state and the desired state, respectively. $R_t > 0$, $Q_t \geq 0$ and $P_T \geq 0$ are penalty weighting matrices on the variations of the input $u(t)$, state $x(t)$ and terminal state x_T from their desired values.

The linear quadratic tracking problem is to find a state feedback control $\{u^*(s), 0 \leq s \leq T\}$ such that J_T is minimized.

6.3 Main Result

We present our main result on the linear quadratic tracking problem. The derivation of the result will be presented in the next section.

First, associated with the system (6.1) and the cost function (6.2), we introduce the following deterministic backward delay system which is needed for computing the optimal tracking control law:

$$\begin{aligned} -\dot{r}(t) = & \Phi'(t)r(t) + Q_t \bar{x}(t) - K(t, t)z(t), \\ z(t) = & R_t \bar{u}(t) + \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) r(t + d_j) \chi_{[t, T]}(t + d_j) \end{aligned} \quad (6.3)$$

$$-\sum_{j=0}^l \int_0^{d_j} \Gamma'_{(j)}(t+d_j)K(t+d_j, t+\tau)z(t+\tau)\chi_{[t,T]}(t+d_j)d\tau, \quad (6.4)$$

where the initial state $r(T) = P_T \bar{x}_T$. By defining $h_t = d \wedge (T - t)$, $K(t + \theta, t)$ can be calculated by

$$K(t + \theta, t) = \sum_{j=0}^l P(t + \theta, t + d_j, t) \Gamma_{(j)}(t + d_j) R_t^{-1} \chi_{[t,T]}(t + d_j), \quad 0 \leq \theta \leq h_t. \quad (6.5)$$

where $P(t + \theta_1, t + \theta_2, t)$, $\theta_1, \theta_2 \in [0, h_t]$ satisfies the following Riccati type partial differential equations:

$$\dot{P}(t, t, t) = -\Phi'(t)P(t, t, t) - P(t, t, t)\Phi(t) - Q_t + K(t, t)R_t K'(t, t), \quad (6.6)$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta} \right) P(t + \theta, t, t) = -P(t + \theta, t, t)\Phi(t) + K(t + \theta, t)R_t K'(t, t), \quad (6.7)$$

$$P(t + \theta_1, t + \theta_2, t) = P'(t + \theta_2, t + \theta_1, t), \quad (6.8)$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta_1} - \frac{\partial}{\partial \theta_2} \right) P(t + \theta_1, t + \theta_2, t) = K(t + \theta_1, t)R_t K'(t + \theta_2, t). \quad (6.9)$$

If we set s as the starting time of system (6.1), then we see that $\tilde{u}_s(t)$, which is defined by

$$\tilde{u}_s(t) \triangleq \sum_{j=1}^l \Gamma_{(j)}(k)u(t - d_j)\chi_{[-d,s]}(t - d_j), \quad t \geq s \geq 0, \quad (6.10)$$

is a combination of the past control inputs. The main result is given below whose proof is deferred to the next section.

Theorem 6.1. *For system (6.1) and quadratic cost function (6.2), the optimal tracking controller $\{u^*(s), 0 \leq s \leq T\}$ that minimizes (6.2) can be calculated by*

$$u^*(s) = R_s^{-1}z(s) - K'(s, s)x(s) - \int_0^{h_s} K'(s + \tau, s)\tilde{u}_s^*(s + \tau)d\tau, \quad (6.11)$$

where $z(s)$ is as defined in the auxiliary backward system (6.3)-(6.4), $K(s + \tau, s)$, $0 \leq \tau \leq h_s$, $\tilde{u}_s^*(k)$ is given in (6.10) with $u(t - d_j)$ replaced by $u^*(t - d_j)$ and $u^*(t) = \mu_t$ for $t < 0$.

6.4 Proof of Main Result

We shall prove the main result by introducing a dual backward stochastic system and establish a duality between the optimal tracking and an estimation problem for the backward stochastic system. Some elementary tools such as projection theory can be used for solving the smoothing problem for the backward stochastic system.

We introduce the notation similar to (3.3)-(3.7). Then $\forall t \geq s$, the system (6.1) can be rewritten as

$$\dot{x}(t) = \Phi(t)x(t) + \Gamma_s(t)u_s(t) + \tilde{u}_s(t), \quad (6.12)$$

and the cost function (6.2) can be split into two terms,

$$J_T = J_T^s + \bar{J}_T^s, \quad (6.13)$$

where

$$\begin{aligned} J_T^s &= (x_T - \bar{x}_T)' P_T (x_T - \bar{x}_T) + \int_s^T (x(\tau) - \bar{x}(\tau))' Q_\tau (x(\tau) - \bar{x}(\tau)) d\tau \\ &\quad + \int_s^T (u_s(\tau) - \bar{u}_s(\tau))' R_s(\tau) (u_s(\tau) - \bar{u}_s(\tau)) d\tau, \end{aligned} \quad (6.14)$$

$$\bar{J}_T^s = \int_0^s [(x(\tau) - \bar{x}(\tau))' Q_\tau (x(\tau) - \bar{x}(\tau)) + (u(\tau) - \bar{u}(\tau))' R_\tau (u(\tau) - \bar{u}(\tau))] d\tau. \quad (6.15)$$

and $\bar{u}_s(t)$ is defined as $u_s(t)$ by replacing $u(t)$ with $\bar{u}(t)$. Obviously, $\{u(t), t \geq s\}$ can be designed to minimize only J_T^s because it has no effect on $\bar{J}_T^s$.

6.4.1 Duality between Tracking and Smoothing Problem

By slightly abusing the notation, we denote $\{x(t), s \leq t \leq T\}$ and $\{u_s(t), s \leq t \leq T\}$ respectively by x_s and u_s . $\bar{x}_s, \bar{u}_s$ are defined in the same way. Since $\tilde{u}_s(t)$ is non-zero only when $s \leq t < s + d$, we denote $\tilde{u}_s$ as $\{\tilde{u}_s(t), s \leq t \leq s + d\}$. From (6.12) we have

$$x(t) = \Psi_{t,s}^\Phi x(s) + \int_s^{s+d} \Psi_{t,\tau}^\Phi \tilde{u}_s(\tau) d\tau + \int_s^t \Psi_{t,\tau}^\Phi \Gamma_s(\tau) u_s(\tau) d\tau \quad (6.16)$$

where $\Psi_{t,\tau}^\Phi$ is the transition matrix generated by $\phi(t)$, i.e. $\frac{d\Psi_{t,\tau}^\Phi}{dt} = \phi(t)\Psi_{t,\tau}^\Phi$. In order to simplify the notation, we drop the explicit dependence on t . Then we write (6.16) in the following compact form

$$x_s = \mathcal{O}_s \xi_s + \mathcal{B}_s u_s \quad (6.17)$$

$$x_T = \mathcal{H}_s \xi_s + \mathcal{C}_s u_s, \quad (6.18)$$

where $\xi_s = \text{col}\{x(s), \tilde{u}_s\}$, the operators $\mathcal{O}_s, \mathcal{B}_s, \mathcal{H}_s, \mathcal{C}_s$ are defined as follows:

$$\mathcal{O}_s : \xi_s \mapsto x_s, x(t) = \Psi_{t,s}^\Phi x(s) + \int_s^{(s+d) \wedge t} \Psi_{t,\tau}^\Phi \tilde{u}_s(\tau) d\tau \quad (6.19)$$

$$\mathcal{B}_s : u_s \mapsto x_s, x(t) = \int_s^t \Psi_{t,\tau}^\Phi \Gamma_s(\tau) u_s(\tau) d\tau \quad (6.20)$$

$$\mathcal{H}_s : \xi_s \mapsto x_T, x_T = \Psi_{T,s}^\Phi x(s) + \int_s^{(s+d) \wedge T} \Psi_{T,\tau}^\Phi \tilde{u}_s(\tau) d\tau \quad (6.21)$$

$$\mathcal{C}_s : u_s \mapsto x_T, x_T = \int_0^T \Psi_{T,\tau}^\Phi \Gamma_s(\tau) u_s(\tau) d\tau, \quad (6.22)$$

We are in the position to introduce the following model which is dual with (6.17)-(6.18):

$$\mathbf{x}_s^0 = \mathcal{H}_s^* \mathbf{x}_T + \mathcal{O}_s^* \mathbf{q}_s, \quad (6.23)$$

$$\mathbf{y}_s = \mathcal{C}_s^* \mathbf{x}_T + \mathcal{B}_s^* \mathbf{v}_s, \quad (6.24)$$

where $\mathbf{x}_s^0 = \text{col}\{\mathbf{x}(s), \mathbf{x}_s^1\}$ with $\mathbf{x}_s^1 \triangleq \{\mathbf{x}(t), s \leq t \leq s+d\}$; $\mathcal{O}_s^*, \mathcal{B}_s^*, \mathcal{H}_s^*, \mathcal{C}_s^*$ are adjoint operators of $\mathcal{O}_s, \mathcal{B}_s, \mathcal{H}_s, \mathcal{C}_s$, respectively; $\mathbf{q}_s, \mathbf{v}_s$ and $\mathbf{y}_s$ represent $\{\mathbf{q}(t), s \leq t \leq T\}$, $\{\mathbf{v}(t), s \leq t \leq T\}$ and $\{\mathbf{y}(t), s \leq t \leq T\}$, respectively, and $\mathbf{x}_T, \mathbf{q}(t), \mathbf{v}(t)$ are uncorrelated white noises with zero means and covariances

$$\langle \mathbf{x}_T, \mathbf{x}_T \rangle = P_T, \quad \langle \mathbf{q}(t_1), \mathbf{q}(t_2) \rangle = Q_{t_1} \delta(t_1 - t_2), \quad \langle \mathbf{v}(t_1), \mathbf{v}(t_2) \rangle = R_s(t_1) \delta(t_1 - t_2).$$

We can represent the model (6.23)-(6.24) by the following backward stochastic system:

$$-\dot{\mathbf{x}}(t) = \Phi'(t)\mathbf{x}(t) + \mathbf{q}(t), \quad \mathbf{x}(T) = \mathbf{x}_T, \quad (6.25)$$

$$\mathbf{y}(t) = \Gamma'_s(t)\mathbf{x}(t) + \mathbf{v}(t), \quad s \leq t \leq T, \quad (6.26)$$

Motivated by the discrete-time setting [72], for matrix value function $M(\tau_1, \tau_2)$, $s \leq \tau_1 \leq T$, $s \leq \tau_2 \leq T$ and vector value function $w(\tau)$, $s \leq \tau \leq T$, we define

$$w' M w \triangleq \int_0^t \int_0^t w'(\tau_1) M(\tau_1, \tau_2) w(\tau_2) d\tau_1 d\tau_2. \quad (6.27)$$

Then we have the following result.

Lemma 6.1. *By making use of the stochastic state-space model (6.25)-(6.26), J_T^s can be put in the following quadratic form*

$$J_T^s = \begin{bmatrix} \xi_s \\ \bar{x}_T \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}' \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_T \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_T \\ -\mathbf{q}_s \\ -\mathbf{v}_s \\ \mathbf{y}_s \end{bmatrix} \right\rangle \begin{bmatrix} \xi_s \\ \bar{x}_T \\ \bar{x}_s \\ \bar{u}_s \\ u_s \end{bmatrix}. \quad (6.28)$$

Proof. The proof is similar to that in [71] and thus omitted here. $\square$

Now, we shall apply a re-organization of measured outputs as that in Chapter 4 to decouple u_s . First, decompose $\mathbf{y}(t)$ and $\mathbf{v}(t)$ of (6.26) as follows:

$$\mathbf{y}(t) = \begin{cases} \text{col} \{ \mathbf{y}_0(t), \dots, \mathbf{y}_i(t) \}, & d_i \leq t - s < d_{i+1}, \\ \text{col} \{ \mathbf{y}_0(t), \dots, \mathbf{y}_l(t) \}, & t - s \geq d, \end{cases} \quad (6.29)$$

$$\mathbf{v}(t) = \begin{cases} \text{col} \{ \mathbf{v}_0(t), \dots, \mathbf{v}_i(t) \}, & d_i \leq t - s < d_{i+1}, \\ \text{col} \{ \mathbf{v}_0(t), \dots, \mathbf{v}_l(t) \}, & t - s \geq d, \end{cases} \quad (6.30)$$

where $\mathbf{y}_i(t)$ and $\mathbf{v}_i(t), i = 0, \dots, l$ satisfy

$$\mathbf{y}_i(t) = \Gamma'_{(i)}(t)\mathbf{x}(t) + \mathbf{v}_i(t), \quad (6.31)$$

with $\langle \mathbf{v}_i(t_1), \mathbf{v}_j(t_2) \rangle = R_{i,t_1} \delta_{i,j} \delta(t_1 - t_2)$. Then we define

$$\bar{\mathbf{y}}(t) = \sum_{j=0}^l \mathbf{y}_j(t + d_j) \chi_{[t,T]}(t + d_j). \quad (6.32)$$

In view of (6.31), $\bar{\mathbf{y}}(t)$ can be rewritten as

$$\bar{\mathbf{y}}(t) = \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \mathbf{x}(t + d_j) \chi_{[t,T]}(t + d_j) + \bar{\mathbf{v}}(t), \quad (6.33)$$

where

$$\bar{\mathbf{v}}(t) = \sum_{j=0}^l \mathbf{v}_j(t + d_j) \chi_{[t,T]}(t + d_j). \quad (6.34)$$

It is easy to verify that $\bar{\mathbf{v}}(t), s \leq t \leq T$ are uncorrelated white noises with zero means and covariances $\langle \bar{\mathbf{v}}(t), \bar{\mathbf{v}}(\tau) \rangle = R_t \delta_{t,\tau}$. We arrive at the following result.

Lemma 6.2. *Let*

$$\begin{aligned} U_s &= \{u(t), s \leq t \leq T\}, \quad \bar{U}_s = \{\bar{u}(t), s \leq t \leq T\}, \\ \bar{\mathbf{y}}_s &= \{\bar{\mathbf{y}}(t), s \leq t \leq T\}, \quad \bar{\mathbf{v}}_s = \{\bar{\mathbf{v}}(t), s \leq t \leq T\}, \end{aligned}$$

where $\bar{\mathbf{y}}(t)$ is as in (6.33) and $\bar{\mathbf{v}}(t)$ in (6.34). Then J_T^s of (6.14) can be further rewritten in the following quadratic form

$$J_T^s = \begin{bmatrix} \xi_s \\ \bar{x}_T \\ \bar{x}_s \\ \bar{U}_s \\ U_s \end{bmatrix}' \left\langle \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_T \\ -\mathbf{q}_s \\ -\bar{\mathbf{v}}_s \\ \bar{\mathbf{y}}_s \end{bmatrix}, \begin{bmatrix} \mathbf{x}_s^0 \\ -\mathbf{x}_T \\ -\mathbf{q}_s \\ -\bar{\mathbf{v}}_s \\ \bar{\mathbf{y}}_s \end{bmatrix} \right\rangle \begin{bmatrix} \xi_s \\ \bar{x}_T \\ \bar{x}_s \\ \bar{U}_s \\ U_s \end{bmatrix}. \quad (6.35)$$

Furthermore, the optimal control is given by

$$U_s^* = R_{\bar{\mathbf{y}}_s}^{-1} (R_{\bar{\mathbf{y}}_s, \mathbf{x}_T} \bar{x}_T + R_{\bar{\mathbf{y}}_s, \mathbf{q}_s} \bar{x}_s + R_{\bar{\mathbf{y}}_s, \bar{\mathbf{v}}_s} \bar{U}_s - R_{\bar{\mathbf{y}}_s, \mathbf{x}_s^0} \xi_s), \quad (6.36)$$

where $U_s^* = \{u^*(t), s \leq t \leq T\}$, $R_{\bar{\mathbf{y}}_s, \mathbf{x}_T} = \langle \bar{\mathbf{y}}_s, \mathbf{x}_T \rangle$, $R_{\bar{\mathbf{y}}_s, \mathbf{q}_s} = \langle \bar{\mathbf{y}}_s, \mathbf{q}_s \rangle$, $R_{\bar{\mathbf{y}}_s, \bar{\mathbf{v}}_s} = \langle \bar{\mathbf{y}}_s, \bar{\mathbf{v}}_s \rangle$, $R_{\bar{\mathbf{y}}_s, \mathbf{x}_s^0} = \langle \bar{\mathbf{y}}_s, \mathbf{x}_s^0 \rangle$, $R_{\bar{\mathbf{y}}_s} = \langle \bar{\mathbf{y}}_s, \bar{\mathbf{y}}_s \rangle$.

Proof. The proof is omitted and a similar proof can be found in [136]. $\square$

Remark 6.1. Note that the feedback gains $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{x}_T}$, $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{q}_s}$, $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \bar{\mathbf{v}}_s}$ and $R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{x}_s^0}$ are the transpose of the optimal estimation gain of $\mathbf{x}_T$, $\mathbf{q}_s$, $\bar{\mathbf{v}}_s$ and $\mathbf{x}_s^0$ based on $\bar{\mathbf{y}}_s$, respectively. Therefore, the problem is converted into the dual problem, i.e. optimal estimation problem, which can be solved by some elementary approach such as projection approach.

6.4.2 Formation of the Backward System and Solution to the Tracking Problem

Lemma 6.2 reveals a duality between the linear quadratic tracking problem and the optimal state and white noise smoothing problem for the stochastic backward system (6.25) and (6.33). We have solved the state estimation problem of system (6.25) and (6.33) in Chapter 4, Theorem 4.1. For easy reference, we recall it here.

Theorem 6.2. For system (6.25) and (6.33), the recursive optimal estimator is given by

$$\dot{\hat{\mathbf{x}}}(t|t) = -\Phi'(t)\hat{\mathbf{x}}(t|t) - K(t, t)\mathbf{e}(t), \quad (6.37)$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial \theta} \right) \hat{\mathbf{x}}(t + \theta|t) = -K(t + \theta, t)\mathbf{e}(t), \quad 0 \leq \theta \leq h_t, \quad (6.38)$$

$$\mathbf{e}(t) = \bar{\mathbf{y}}(t) - \sum_{j=0}^l [\Gamma'_{(j)}(t + d_j)\hat{\mathbf{x}}(t + d_j|t)\chi_{[t, T]}(t + d_j)], \quad (6.39)$$

where $\mathbf{e}(t)$ is the innovation sequence with covariance R_t . $K(t + \theta|t)$, $0 \leq \theta \leq h_t$ can be calculated by (6.5). The initial condition is:

$$\hat{\mathbf{x}}(T|T) = 0. \quad (6.40)$$

Theorem 6.2 provides the optimal estimate of $\mathbf{x}_s^0$. However, the gain $R_{\mathbf{x}_s^0, \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1}$ is not so explicit. Moreover, the white noise estimate is not straightforward. The following theorem is needed based on which the above-mentioned problems become transparent.

Theorem 6.3. For system (6.25) and (6.33), $\mathbf{e}(t)$, defined in Theorem 6.2, satisfies

$$-\dot{\tilde{\mathbf{x}}}(t) = \Phi'(t)\tilde{\mathbf{x}}(t) + \mathbf{q}(t) - K(t, t)\mathbf{e}(t), \quad (6.41)$$

$$\begin{aligned} \mathbf{e}(t) = & \bar{\mathbf{v}}(t) + \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \tilde{\mathbf{x}}(t + d_j | t + d_j) \chi_{[t, N]}(t + d_j) \\ & - \sum_{j=0}^l \int_0^{d_j} \Gamma'_{(j)}(t + d_j) K(t + d_j, t + \tau) \mathbf{e}(t + \tau) \chi_{[t, N]}(t + d_j) d\tau \end{aligned} \quad (6.42)$$

with the initial condition $\tilde{\mathbf{x}}(T|T) = \mathbf{x}_T$.

Proof. Subtracting (6.37) from (6.25) yields (6.41). By applying projection theory, we have

$$\begin{aligned} \hat{\mathbf{x}}(t + \theta|t) &= \int_t^T \langle \mathbf{x}(t + \theta), \mathbf{e}(\tau) \rangle R_\tau^{-1} \mathbf{e}(\tau) d\tau \\ &= \hat{\mathbf{x}}(t + \theta|t + \theta) + \int_0^\theta \langle \mathbf{x}(t + \theta), \mathbf{e}(t + \tau) \rangle R_{t+\tau}^{-1} \mathbf{e}(t + \tau) d\tau \\ &= \hat{\mathbf{x}}(t + \theta|t + \theta) + \int_0^\theta \sum_{j=0}^l P(t + \theta, t + \tau + d_j, t + \tau) \Gamma_{(j)}(t + \tau + d_j) R_{t+\tau}^{-1} \\ &\quad \times \chi_{[t+\tau, T]}(t + \tau + d_j) \mathbf{e}(t + \tau) d\tau \\ &= \hat{\mathbf{x}}(t + \theta|t + \theta) + \int_0^\theta K(t + \theta, t + \tau) \mathbf{e}(t + \tau) d\tau. \end{aligned} \quad (6.43)$$

Subtracting (6.43) from $\mathbf{x}(t + \theta)$ yields

$$\tilde{\mathbf{x}}(t + \theta|t) = \tilde{\mathbf{x}}(t + \theta|t + \theta) - \int_0^\theta K(t + \theta, t + \tau) \mathbf{e}(t + \tau) d\tau. \quad (6.44)$$

Since

$$\mathbf{e}(t) = \bar{\mathbf{y}}(t) - \hat{\mathbf{y}}(t|t) = \sum_{j=0}^l \Gamma'_{(j)}(t + d_j) \tilde{\mathbf{x}}(t + d_j|t) \chi_{[t, T]}(t + d_j) + \bar{\mathbf{v}}(t). \quad (6.45)$$

(6.42) is obtained by substituting (6.44) into (6.45). $\square$

Note that system (6.42)-(6.41) is linear. Therefore, $\mathbf{e}(s)$ can be written in the following way

$$\mathbf{e}(s) = \Psi_{T,s}^x \mathbf{x}_T + \int_s^T \Psi_{\tau,s}^q \mathbf{q}(\tau) d\tau + \int_s^T \Psi_{\tau,s}^v \bar{\mathbf{v}}(\tau) d\tau, \quad (6.46)$$

where $\Psi_{T,s}^x$, $\Psi_{\tau,s}^q$ and $\Psi_{\tau,s}^v$ are transition matrices from $\mathbf{x}_T$, $\mathbf{q}(\tau)$ and $\bar{\mathbf{v}}(\tau)$ to $\mathbf{e}(s)$, respectively. Due to the independence of $\mathbf{x}_T$, $\mathbf{q}(\tau)$ and $\bar{\mathbf{v}}(\tau)$, we have

$$R_{\mathbf{e}(s), \mathbf{x}_T} \triangleq \langle \mathbf{e}(s), \mathbf{x}_T \rangle = \Psi_{T,s}^x P_T, \quad (6.47)$$

$$R_{\mathbf{e}(s), \mathbf{q}(t)} \triangleq \langle \mathbf{e}(s), \mathbf{q}(t) \rangle = \Psi_{t,s}^q Q_t, \quad t > s, \quad (6.48)$$

$$R_{\mathbf{e}(s), \bar{\mathbf{v}}(t)} \triangleq \langle \mathbf{e}(s), \bar{\mathbf{v}}(t) \rangle = \Psi_{t,s}^v R_t, \quad t \geq s. \quad (6.49)$$

Define a linear operator $\mathcal{G} : L_2(s, T; \mathbb{R}^n) \mapsto L_2(s, T; \mathbb{R}^n)$ by $\mathbf{e}_s = \mathcal{G} \bar{\mathbf{y}}_s$, where $\mathbf{e}_s = \{\mathbf{e}(t), s \leq t \leq T\}$, and Dirac distribution function $T_{\delta_r} : L_2(s, T; \mathbb{R}^n) \mapsto \mathbb{R}^n$ which is a generalized function concentrated at the point r , i.e. $T_{\delta_r} x = x(r)$, $\forall x \in L_2(s, T; \mathbb{R}^n)$. According to the relationship between $\mathbf{e}_s$ and $\bar{\mathbf{y}}_s$, we have

$$T_{\delta_0} \mathcal{G}^* x = T_{\delta_0} x = x(0), \quad \forall x \in L_2(s, T; \mathbb{R}^n). \quad (6.50)$$

By projection theory, we have

$$R_{\mathbf{x}_T, \bar{\mathbf{y}}_s} R_{\bar{\mathbf{y}}_s}^{-1} \bar{\mathbf{y}}_s = \int_s^T R_{\mathbf{x}_T, \mathbf{e}(t)} R_t^{-1} \mathbf{e}(t) dt \triangleq \mathcal{S}_x \mathbf{e}_s. \quad (6.51)$$

By combining (6.47), (6.50) and (6.51) we get

$$T_{\delta_0} R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{x}_T} = T_{\delta_0} \mathcal{G}^* \mathcal{S}_x^* = R_s^{-1} \Psi_{T,s}^x P_T. \quad (6.52)$$

Similarly we can get

$$\begin{aligned} T_{\delta_0} R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{q}(t)} &= R_s^{-1} \Psi_{t,s}^q Q_t, \quad T_{\delta_0} R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \bar{\mathbf{v}}(t)} = R_s^{-1} \Psi_{t,s}^v R_t, \\ T_{\delta_0} R_{\bar{\mathbf{y}}_s}^{-1} R_{\bar{\mathbf{y}}_s, \mathbf{x}(s+\tau)} &= K'(s+\tau, s). \end{aligned}$$

According to (6.36) in Lemma 6.2, the optimal input at time instant s is given as follow

$$\begin{aligned} u^*(s) &= T_{\delta_0} U_s^* \\ &= R_s^{-1} \left[\Psi_{T,s}^x P_T \bar{x}_T + \int_s^T \Psi_{t,s}^q Q_t \bar{x}(t) dt + \int_s^T \Psi_{t,s}^v R_t \bar{u}(t) dt \right] \\ &\quad - K'(s, s) x(s) - \int_0^{h_s} K'(s+\tau, s) \tilde{u}_s^*(s+\tau) d\tau. \end{aligned} \quad (6.53)$$

By comparing (6.46) and the first term of (6.53), it is noted that if we replace $\mathbf{x}_T$, $\mathbf{q}(t)$ and $\bar{\mathbf{v}}(t)$ with $P_T \bar{x}_T$, $Q_t \bar{x}(t)$ and $R_t \bar{u}(t)$, respectively, the system becomes deterministic system (6.3)-(6.4) and the optimal feedback controller in (6.11) is obtained.

6.5 Conclusion

In this chapter, we have investigated the linear quadratic tracking problem for linear continuous-time systems with multiple delays in a single input channel. By establishing the duality between the tracking problem and the white noise smoothing

problem for a backward stochastic system, we are able to solve the dual problem instead. The optimal input is given in terms of the solution of a Riccati type partial differential equation and the output of an auxiliary backward delay system which is formulated for the first time.

Part II

Distributed Control of Multi-agent Systems with Communication/Input Delays

Chapter 7

Distributed Consensus for MAS with Delayed Control Input Information from Neighbors

In this chapter, we consider the consensus of continuous MASs with single-integrator dynamics for each node. We comparatively study two control protocols: (a) a standard local voting protocol; (b) a protocol that includes information about the control inputs of neighbors. We first demonstrate the advantages and better performance by exploiting the control input information from neighbors. Specifically, for a certain quadratic performance index, the standard protocol (a) yields an optimal cost that depends on the graph topology, while protocol (b) yields an optimal value that is independent of the topology. Moreover, the minimum value of the cost under protocol (a) is not less than that under protocol (b). What is more, under protocol (b), the convergence rate can be selected arbitrarily, independently of the topology. This is not the case for protocol (a).

On the other hand, from implementation point of view, the control input information from neighbors can only be received after a time delay. Therefore, a distributed protocol based on neighbors' delayed control input information is proposed and investigated. It is shown that the proposed protocol can guarantee the consensus of

the MAS for any nonnegative delay. Furthermore, the convergence rate in relation to the delay is obtained. When the time delay tends to 0, the convergence rate approaches that under delay-free case, which is superior to the performance of the standard local voting protocol. An application of the proposed protocol to formation control is given thereafter. Some numerical examples and experiments demonstrate the effectiveness of the results.

7.1 Problem Statement

We shall formulate the consensus problem to be studied for MASs. The system to be considered consists of n autonomous agents. The communication topology can be represented by a graph $\mathcal{G}$ (see Appendix A for graph theory notations). If we label the agents from 1 to n , we can see that $\mathcal{V} = \{1, \dots, n\}$. The agent i is assumed to have the following dynamics

$$\dot{x}_i(t) = u_i(t), \quad i \in \mathcal{V}, \quad (7.1)$$

where $x_i \in \mathbb{R}$ and $u_i \in \mathbb{R}$ are the state and control input of agent i , respectively. It should be noted that the results to be developed later can be extended to the vector state case by introducing Kronecker product which, however, is omitted here for simplicity.

We say that the MAS achieves consensus if

$$\lim_{t \rightarrow \infty} (x_i(t) - x^o(t)) = 0, \quad i = 1, 2, \dots, n, \quad (7.2)$$

where $x^o(t)$ is the reference state. Our objective is to design a distributed protocol such that (7.2) holds. In most networks, $x^o(t)$ may not be available to all the agents. We denote $\bar{\mathcal{G}}$ the graph by incorporating the $(n + 1)$ -th vertex corresponding to $x^o(t)$ into $\mathcal{G}$. Correspondingly, we define $\{\bar{\mathcal{V}}, \bar{\mathcal{E}}, \bar{A}\}$ the triplet of $\bar{\mathcal{G}}$. Since $x^o(t)$ is not affected by all the agents, the $(n + 1)$ -th vertex can be regarded as a virtual

leader. Hence the last row of $\bar{A}$ is 0. It is straightforward that A is the n -th leading principle minor of $\bar{A}$, therefore we have

$$\bar{A} = \left[\begin{array}{ccc|c} & & & b_1 \\ & A & & \vdots \\ & & & b_n \\ \hline 0 & \cdots & 0 & 0 \end{array} \right], \quad (7.3)$$

where b_i indicates the accessibility of $x^o(t)$ by agent i . When $x^o(t)$ is accessible by agent i , $b_i > 0$, otherwise $b_i = 0$. We use D , L and $\bar{D}$, $\bar{L}$ to denote the in-degree matrices and Laplacian matrices of $\mathcal{G}$ and $\bar{\mathcal{G}}$, respectively. Let

$$b = \text{col}\{b_1, b_2, \dots, b_n\}, \quad B = \text{diag}\{b_1, b_2, \dots, b_n\},$$

then we have

$$\bar{L} = \left[\begin{array}{ccc|c} L + B & & & -b \\ \hline 0 & \cdots & 0 & 0 \end{array} \right]. \quad (7.4)$$

The following two protocols, which can be found in [99], are proposed:

$$u_{1,i}(t) = -\frac{\beta}{d_i + b_i} \left[\sum_{j \in \mathcal{V}} a_{i,j}(x_i(t) - x_j(t)) + b_i(x_i(t) - x^o(t)) \right], \quad (7.5)$$

$$u_{2,i}(t) = \frac{1}{d_i + b_i} \left\{ -\gamma \left[\sum_{j \in \mathcal{V}} a_{i,j}(x_i(t) - x_j(t)) + b_i(x_i(t) - x^o(t)) \right] + \sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t) \right\}, \quad (7.6)$$

where $\beta > 0$, $\gamma > 0$ are parameters to be determined. Note that the two protocols (7.5) and (7.6) are standard and widely studied by the researchers, see [51], [86], [99], etc. We comparatively study the two protocols and show that protocol (7.6) yields a better performance than protocol (7.5) does.

Before proceeding, the following result is introduced which can be found in [51] and [99]. For the convenience of developing our result, we provide the proof of the following result using a different method.

Lemma 7.1. *Given that $x^o(t) \equiv x^o$ is a constant, for MAS (7.1) with undirected communication graph $\mathcal{G}$, if the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree, the objective (7.2) can be achieved by $u_i(t) = u_{1,i}(t)$ or $u_i(t) = u_{2,i}(t)$.*

Proof. Define

$$x(t) = \text{col}\{x_1(t), x_2(t), \dots, x_n(t)\}, \quad \tilde{x}(t) = \text{col}\{\tilde{x}_1(t), \tilde{x}_2(t), \dots, \tilde{x}_n(t)\}, \quad (7.7)$$

with $\tilde{x}_i(t) = x_i(t) - x^o$. Since graph $\bar{\mathcal{G}}$ contains a spanning tree, we have $d_i + b_i \neq 0$, which makes the protocols well defined. By using the fact that $L\mathbf{1}_n = 0$ and $B\mathbf{1}_n = b$, we rewrite protocol (7.5) as follows

$$u_{(1)}(t) = -\beta(D + B)^{-1}[(L + B)x(t) - bx^o] = -\beta(D + B)^{-1}(L + B)\tilde{x}(t),$$

where

$$u_{(1)}(t) = \text{col}\{u_{1,1}(t), \dots, u_{1,n}(t)\}. \quad (7.8)$$

The closed-loop system is

$$\dot{\tilde{x}}(t) = -\beta(D + B)^{-1}(L + B)\tilde{x}(t). \quad (7.9)$$

It is noted that $L + B > 0$. In fact, if there exists 0 eigenvalue in $L + B$, there are more than one 0 eigenvalues in $\bar{L}$ which, according to Lemma A.1, contradicts the assumption that the augmented graph $\bar{\mathcal{G}}$ contains a spanning tree. Since $D + B > 0$, according to [43], we know that all the eigenvalues of $(D + B)^{-1}(L + B)$ are positive real numbers. Therefore, system (7.9) is stable for all $\beta > 0$. According to the definition of $\tilde{x}(t)$, (7.2) is obtained.

Now consider protocol (7.6). Define

$$u_{(2)}(t) = \text{col}\{u_{2,1}(t), u_{2,2}(t), \dots, u_{2,n}(t)\}. \quad (7.10)$$

Then according to (7.6) we have

$$\begin{aligned} u_{(2)}(t) &= (D + B)^{-1} \{A u_{(2)}(t) - \gamma[(L + B)x(t) - bx^o]\} \\ &= (L + B)^{-1}(-\gamma)[(L + B)x(t) - bx^o] \\ &= -\gamma \tilde{x}(t), \end{aligned}$$

and the closed-loop system is

$$\dot{\tilde{x}}(t) = -\gamma \tilde{x}(t). \quad (7.11)$$

It is obvious that system (7.11) is stable for all $\gamma > 0$, which completes the proof. $\square$

There are some advantages in the protocol (7.6), which will be shown in the next section. However, (7.6) cannot be implemented due to the extra term $\sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t)$. In fact, $\forall (i, j) \in \mathcal{E}$, the control input $u_{2,i}(t)$ depends on $u_{2,j}(t) = \dot{x}_j(t)$ and vice versa. From the implementation point of view, we introduce a small delay τ in the extra term $\sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t)$. The problems in this chapter are that

1. What is the condition on τ under which the consensus is guaranteed;
2. How does τ affect the consensus speed.

7.2 Main Results

We shall first compare the two protocols, and analyze the second protocol with an introduced delay.

7.2.1 Comparison

From the proof of Lemma 7.1 we can see that both the two protocols can guarantee the consensus and an arbitrary convergence rate can be achieved by tuning β in (7.5) or γ in (7.6). However, in protocol (7.5), the convergence rate is also affected by the smallest eigenvalue of $(D + B)^{-1}(L + B)$ in view of (7.9), which may not

be available to every agent. Furthermore, when $\dot{x}^o(t) \neq 0$, protocol (7.5) fails to synchronize the network [98]. While, by slightly modifying the protocol (7.6), the consensus can still be guaranteed. This will be shown later.

Next, we will show that for a given quadratic cost function, we can choose β and γ to minimize it. Moreover, the minimum value by applying protocol (7.6) will not exceed the one by applying (7.5). To this end, denote $0 < \eta_1 \leq \eta_2 \leq \dots \leq \eta_n$ the n eigenvalues of $(D + B)^{-1}(L + B)$ and define the cost function

$$J = \int_0^\infty \sum_{i=1}^n (d_i + b_i) (\kappa_1 \tilde{x}_i^2(t) + \kappa_2 u_i^2(t)) dt, \quad (7.12)$$

where $\kappa_1 > 0$, $\kappa_2 > 0$ are positive numbers, u_i is chosen to be either $u_{1,i}(t)$ or $u_{2,i}(t)$. $d_i + b_i$ is the in-degree of agent i which reflects the significance of agent i in the network. Therefore, we choose $d_i + b_i$ as the weight of the quadratic cost of agent i in (7.12). The relative difference of weights over state and control input is indicated by κ_1 and κ_2 . The following lemma is introduced before providing the result.

Lemma 7.2. *Define $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\} > 0$. $F(\Lambda, y)$ is minimized for all $y \in \mathbb{R}^n$ when $\lambda_1 = \dots = \lambda_n$, where $F(\Lambda, y) = (y' \Lambda^{-1} y)(y' \Lambda y)$. Moreover, the minimum value is $F^* = (y' y)^2$.*

Proof. We write $F(\Lambda, y) = (\sum_{i=1}^n y_i^2 \lambda_i^{-1}) (\sum_{i=1}^n y_i^2 \lambda_i)$, where y_i is the i -th component of y . Without loss of generality, we assume that $\forall i = 1, \dots, n$, $y_i \neq 0$. First, we consider $n = 2$. Then

$$F(\Lambda, y) = (y_1^2 \lambda_1^{-1} + y_2^2 \lambda_2^{-1}) (y_1^2 \lambda_1 + y_2^2 \lambda_2) = y_1^4 + y_2^4 + y_1^2 y_2^2 (\lambda_1^{-1} \lambda_2 + \lambda_2^{-1} \lambda_1).$$

It is obvious that when $\lambda_1 = \lambda_2$, $F(\Lambda, y)$ is minimized for all $y \in \mathbb{R}^2$. Next, assume that this holds for $n = k - 1 \geq 2$. When $n = k$, $F(\Lambda, y)$ can be written as

$$F(\Lambda, y) = \left(y_k^2 \lambda_k^{-1} + \sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1} \right) \left(y_k^2 \lambda_k + \sum_{i=1}^{k-1} y_i^2 \lambda_i \right)$$

$$= y_k^4 + \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1} \right) \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i \right) + y_k^2 \lambda_k \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1} \right) + y_k^2 \lambda_k^{-1} \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i \right).$$

Consider the above equation as a function of λ_k . Note that when

$$\lambda_k = \lambda_k^* = \sqrt{\left(\sum_{i=1}^{k-1} y_i^2 \lambda_i \right) \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1} \right)^{-1}}, \quad (7.13)$$

$F(\Lambda, y)$ is minimized and the minimized value is

$$\min_{\lambda_k} F(\Lambda, y) = \left(y_k^2 + \sqrt{\left(\sum_{i=1}^{k-1} y_i^2 \lambda_i \right) \left(\sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1} \right)} \right)^2.$$

On the other hand, we know that $(\sum_{i=1}^{k-1} y_i^2 \lambda_i)(\sum_{i=1}^{k-1} y_i^2 \lambda_i^{-1})$ is minimized when $\lambda_1 = \dots = \lambda_{k-1} = \alpha$, so

$$\min_{\lambda_1, \dots, \lambda_{k-1}} \left\{ \min_{\lambda_k} F(\Lambda, y) \right\} = \left(\sum_{i=1}^k y_i^2 \right)^2,$$

and according to (7.13), we have

$$\lambda_k^* = \sqrt{\left(\sum_{i=1}^{k-1} y_i^2 \alpha \right) \left(\sum_{i=1}^{k-1} y_i^2 \alpha^{-1} \right)^{-1}} = \alpha,$$

which implies that $F(\Lambda, y)$ is minimized when $\lambda_1 = \dots = \lambda_k$. By induction, the result is proved. $\square$

Theorem 7.1. *Consider MAS (7.1) with undirected communication graph $\mathcal{G}$ and $\dot{x}^o = 0$ in (7.2). Assume that the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree. Then we have the following statements:*

1) *Under protocol (7.5), $J < \infty$ for any $\beta > 0$. Further, when*

$$\beta = \sqrt{\frac{\kappa_1 \tilde{x}'(0)(D+B)(L+B)^{-1}(D+B)\tilde{x}(0)}{\kappa_2 \tilde{x}'(0)(L+B)\tilde{x}(0)}}, \quad (7.14)$$

J is minimized and the minimum value of J is given by

$$J_1 = \sqrt{\kappa_1 \kappa_2 \tilde{x}'(0)(D+B)(L+B)^{-1}(D+B)\tilde{x}(0)} \cdot \sqrt{\tilde{x}'(0)(L+B)\tilde{x}(0)}. \quad (7.15)$$

2) Under protocol (7.6), $J < \infty$ for any $\gamma > 0$. Furthermore, when $\gamma = \sqrt{\kappa_1/\kappa_2}$, J is minimized, and the minimum J is given by

$$J_2 = \sqrt{\kappa_1 \kappa_2} \tilde{x}'(0)(D+B)\tilde{x}(0).$$

3) $J_1 \geq J_2$.

Proof. We first consider protocol (7.5). Since $(D+B)^{-1/2}(L+B)(D+B)^{-1/2}$ is a positive definite matrix, we have the following decomposition

$$(D+B)^{-1/2}(L+B)(D+B)^{-1/2} = P'\eta P, \quad (7.16)$$

where $\eta = \text{diag}\{\eta_1, \dots, \eta_n\}$, P is an orthogonal matrix. Define

$$\bar{x}(t) = P(D+B)^{1/2}\tilde{x}(t). \quad (7.17)$$

then we arrive at that

$$\dot{\bar{x}}_i(t) = -\beta \eta_i \bar{x}_i(t), \quad (7.18)$$

where $\bar{x}_i(t)$ is the i -th component of $\bar{x}(t)$. From (7.18) we obtain that

$$\bar{x}_i(t) = e^{-\beta \eta_i t} \bar{x}_i(0). \quad (7.19)$$

Then we have

$$\begin{aligned} J(\beta) &= \int_0^\infty (\kappa_1 \tilde{x}'(t)(D+B)\tilde{x}(t) + \kappa_2 u'_{(1)}(t)(D+B)u_{(1)}(t)) dt \\ &= \int_0^\infty (\kappa_1 \bar{x}'(t)\bar{x}(t) + \kappa_2 \dot{\bar{x}}'(t)\dot{\bar{x}}(t)) dt \end{aligned}$$

$$= \frac{\kappa_1}{2\beta} \sum_{i=1}^n \frac{\bar{x}_i^2(0)}{\eta_i} + \frac{\kappa_2\beta}{2} \sum_{i=1}^n \bar{x}_i^2(0)\eta_i, \quad (7.20)$$

where $u_{(1)}$ is as defined in (7.8). Obviously, for any $\beta > 0$, $J(\beta)$ is bounded if the disagreement of the initial state of each agent from x^o is bounded. From (7.20) we know that the optimal β is given by

$$\begin{aligned} \beta &= \sqrt{\left(\sum_{i=1}^n \kappa_1 \bar{x}_i^2(0) \eta_i^{-1} \right) \left(\sum_{i=1}^n \kappa_2 \bar{x}_i^2(0) \eta_i \right)^{-1}} \\ &= \sqrt{(\kappa_1 \bar{x}'(0) \eta^{-1} \bar{x}(0)) (\kappa_2 \bar{x}'(0) \eta \bar{x}(0))^{-1}} \\ &= \sqrt{\frac{\kappa_1 \tilde{x}'(0) (D+B) (L+B)^{-1} (D+B) \tilde{x}(0)}{\kappa_2 \tilde{x}'(0) (L+B) \tilde{x}(0)}}, \end{aligned}$$

which is (7.14). The last equality comes from (7.16). The minimum value of $J(\beta)$ is

$$\begin{aligned} J_1 &= \sqrt{\left(\sum_{i=1}^n \kappa_1 \bar{x}_i^2(0) \eta_i^{-1} \right) \left(\sum_{i=1}^n \kappa_2 \bar{x}_i^2(0) \eta_i \right)} \\ &= \sqrt{(\kappa_1 \bar{x}'(0) \eta^{-1} \bar{x}(0)) (\kappa_2 \bar{x}'(0) \eta \bar{x}(0))} \\ &= \sqrt{\kappa_1 \kappa_2 \tilde{x}'(0) (D+B) (L+B)^{-1} (D+B) \tilde{x}(0)} \cdot \sqrt{\tilde{x}'(0) (L+B) \tilde{x}(0)}, \end{aligned}$$

which is (7.15).

Next, we consider protocol (7.6). According to (7.11) we have $\tilde{x}(t) = e^{-\gamma t} \tilde{x}(0)$ and $u_{(2)}(t) = -\gamma e^{-\gamma t} \tilde{x}(0)$, where $u_{(2)}$ is as defined in (7.10). Hence

$$\begin{aligned} J(\gamma) &= \int_0^\infty (\kappa_1 e^{-2\gamma t} + \kappa_2 \gamma^2 e^{-2\gamma t}) \tilde{x}'(0) (D+B) \tilde{x}(0) dt \\ &= \frac{\kappa_1 + \kappa_2 \gamma^2}{2\gamma} \tilde{x}'(0) (D+B) \tilde{x}(0). \end{aligned}$$

It is straightforward that $J(\gamma) < \infty$ for any $\gamma > 0$. Furthermore,

$$J_2 = J \Big|_{\gamma=\sqrt{\frac{\kappa_1}{\kappa_2}}} = \sqrt{\kappa_1 \kappa_2} \tilde{x}'(0) (D+B) \tilde{x}(0).$$

Thus, 2) is proved.

Finally, according to Lemma 7.2, when $\eta_1 = \eta_2 = \dots = \eta_n$, J_1 is minimized and

$$\min J_1 = \sqrt{\kappa_1 \kappa_2} \tilde{x}'(0) \tilde{x}(0) = \sqrt{\kappa_1 \kappa_2} \tilde{x}'(0) (D + B) \tilde{x}(0) = J_2. \quad (7.21)$$

□

Remark 7.1. *From Theorem 7.1 we can see that under the quadratic cost (7.12), the optimal controller gain β of protocol (7.5) cannot be designed without the knowledge of graph, which, however, is not the case for the protocol (7.6). Moreover, the minimum value of the cost under protocol (7.5) is no less than that under protocol (7.6).*

It will now be shown that protocol (7.6) with information of neighbors' control inputs allows arbitrarily fast convergence to consensus independent of the graph topology as long as the augmented graph contains a directed spanning tree, while protocol (7.5) has convergence rate that depends strongly on the graph topology.

Theorem 7.2. *Consider MAS (7.1) with undirected communication graph $\mathcal{G}$ and $\dot{x}^o = 0$ in (7.2). Assume that the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree. Given the proposed two protocols (7.5) and (7.6) with the same control gain, i.e. $\beta = \gamma \triangleq \varrho$, we have the following statements:*

1) Under protocol (7.5),

$$|x_i(t) - x^o| \leq M e^{-\eta_1 \varrho t}, \quad \forall i = 1, \dots, n;$$

2) Under protocol (7.6),

$$|x_i(t) - x^o| \leq M e^{-\varrho t}, \quad \forall i = 1, \dots, n;$$

where $M > 0$ is a constant, η_1 is the smallest eigenvalue of $(D + B)^{-1}(L + B)$.

Proof. The results can be established easily in view of the proofs of Lemma 7.1 and Theorem 7.1. □

Remark 7.2. *The consensus protocol (7.5) has convergence rate $O(e^{-\eta_1 t})$ which depends on the graph topology, including both the structure of the original graph through L and D , and the selection of the pinning nodes through B . Moreover, it can be proved that $0 < \eta_1 \leq 1$, and the equality holds if and only if $A = 0$. This will be discussed later. By contrast, the convergence rate of protocol (7.6) is independent of the graph topology and can be set to be arbitrarily fast through a proper choice of the control gain γ .*

The above clearly demonstrates the advantages and better performance of protocol (7.6). However, protocol (7.6) relies on the instantaneous control input of each neighbor which may not be available. Note that the instantaneous relative position may be obtained by a ranging sensor of agent i . Next, we will investigate the implementation issue of protocol (7.6) by introducing a delay in the protocol.

7.2.2 Protocol with Delayed Control Input from Neighbors

Since the control inputs of each pair of neighbors are coupled, the protocol (7.6) cannot be directly implemented. We shall introduce a small delay τ in the neighbors' control inputs. To be specific, we provide the following protocol

$$u_i(t) = \frac{1}{d_i + b_i} \left\{ -\gamma \left[\sum_{j \in \mathcal{V}} a_{i,j} (x_i(t) - x_j(t)) + b_i (x_i(t) - x^o) \right] + \sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t - \tau) \right\}. \quad (7.22)$$

Before providing the main result, the following lemma from [125] is recalled.

Lemma 7.3. *The linear neutral differential system*

$$\dot{x}(t) - \sum_{k=1}^m B_k \dot{x}(t - k\tau) = \sum_{k=0}^m A_k x(t - k\tau),$$

is stable, independent of delay, if and only if

1)

$$\det \left[I - \sum_{k=1}^m B_k e^{\iota k \theta} \right] \neq 0, \quad \forall \theta \in [0, 2\pi],$$

2)

$$\operatorname{Re} \lambda \left[\left(I - \sum_{k=1}^m B_k \right)^{-1} \sum_{k=0}^m A_k \right] < 0,$$

3)

$$\det \left[\iota y \left(I - \sum_{k=1}^m B_k e^{\iota k \theta} \right) - \sum_{k=0}^m A_k e^{\iota k \theta} \right] \neq 0, \quad \forall \theta \in [0, 2\pi], \quad \forall y \in \mathbb{R}, \quad y \neq 0,$$

where $\iota = \sqrt{-1}$.

It is noted that the matrix $(D + B)^{-1}A$ is similar to $(D + B)^{-1/2}A(D + B)^{-1/2}$ which is a symmetric matrix. Therefore, all the eigenvalues of $(D + B)^{-1}A$ are real numbers. Denote the eigenvalues by $\mu_1 \leq \mu_2 \leq \dots \leq \mu_n$. Then we have the following result.

Theorem 7.3. *Consider MAS (7.1) with undirected communication graph $\mathcal{G}$. If the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree and $\dot{x}^o(t) = 0$, the whole network is synchronized under the protocol (7.22) for any $\tau \geq 0$.*

Proof. We rewrite the whole system in the following way:

$$\begin{cases} \dot{x}(t) = u(t), \\ u(t) = (D + B)^{-1} \{Au(t - \tau) - \gamma [(L + B)x(t) - bx^o]\}, \end{cases} \quad (7.23)$$

where $u(t) = \operatorname{col}\{u_1(t), u_2(t), \dots, u_n(t)\}$. By considering the first equation, the second equation can be written as the following neutral delay system

$$\dot{x}(t) - (D + B)^{-1}A\dot{x}(t - \tau) = -\gamma(D + B)^{-1}[(L + B)x(t) - bx^o].$$

By considering (7.7), we have

$$\dot{\tilde{x}}(t) - (D + B)^{-1}A\dot{\tilde{x}}(t - \tau) = -\gamma(D + B)^{-1}(L + B)\tilde{x}(t). \quad (7.24)$$

It is obvious that the consensus problem is tantamount to the stability problem of system (7.24).

If we set $m = 1$, $A_0 = -\gamma(D + B)^{-1}(L + B)$, $A_1 = 0$, $B_1 = -(D + B)^{-1}A$ in Lemma 7.3, we can get the following conditions which guarantee the stability of system (7.24):

1. $\det [I - B_1 e^{i\theta}] \neq 0, \forall \theta \in [0, 2\pi]$,
2. $\operatorname{Re} \lambda [(I - B_1)^{-1}A_0] < 0$,
3. $\det [\iota y(I - B_1 e^{i\theta}) - A_0] \neq 0, \forall y \in \mathbb{R}/\{0\}, \forall \theta \in [0, 2\pi]$.

Condition 1 is satisfied if all the eigenvalues of B_1 are inside the unit circle. This can be proved by considering the following scaled adjacency matrix

$$A_{scale} = \begin{bmatrix} D + B & 0 \\ 0 & I \end{bmatrix}^{-1} \bar{A} = \begin{bmatrix} (D + B)^{-1}A & (D + B)^{-1}b \\ 0 & 0 \end{bmatrix},$$

and corresponding Laplacian matrix

$$L_{scale} = \begin{bmatrix} I - (D + B)^{-1}A & -(D + B)^{-1}b \\ 0 & 0 \end{bmatrix}.$$

We call A_{scale} the scaled adjacency matrix since $D + B$ is a diagonal matrix and A_{scale} is obtained by scaling the rows of $\bar{A}$ defined in (7.3) such that each row sum of A_{scale} is 1 except the last row. Since graph $\bar{\mathcal{G}}$ contains a spanning tree, the following matrix

$$I - L_{scale} = \begin{bmatrix} (D + B)^{-1}A & (D + B)^{-1}b \\ 0 & 1 \end{bmatrix}$$

is an SIA (stochastic indecomposable and aperiodic) matrix [127]. According to [51], we know that all of the eigenvalues of $(D + B)^{-1}A$ strictly lie inside the unit disk, i.e.

$$-1 < \mu_1 \leq \cdots \leq \mu_n < 1. \quad (7.25)$$

So, the first condition is satisfied.

$$\begin{aligned}
(I - B_1)^{-1}A_0 &= [I - (D + B)^{-1}A]^{-1}[-\gamma(D + B)^{-1}(D + B - A)] \\
&= [L + B]^{-1}[-\gamma(L + B)] \\
&= -\gamma I,
\end{aligned}$$

which meets the second condition.

$$\begin{aligned}
\iota y(I - B_1 e^{i\theta}) - A_0 &= \iota y[I - (D + B)^{-1}Ae^{i\theta}] + \gamma(D + B)^{-1}(D + B - A) \\
&= (D + B)^{-1} \{ \iota y[D + B - A(\cos \theta + i \sin \theta)] + \gamma(L + B) \} \\
&= (D + B)^{-1} \{ (\iota y + \gamma)(D + B) \\
&\quad - (\iota y \cos \theta - y \sin \theta + \gamma)A \} \\
&= (\iota y + \gamma)I - (\iota y \cos \theta - y \sin \theta + \gamma)(D + B)^{-1}A.
\end{aligned}$$

Then condition 3 is equivalent to

$$(\iota y + \gamma) - \mu_i(\iota y \cos \theta - y \sin \theta + \gamma) \neq 0, \forall i = 1, \dots, n. \quad (7.26)$$

According to (7.25), we can see that the imaginary component in (7.26) cannot go to 0, which verifies the third condition. Therefore, the system (7.24) is stable and the consensus can be achieved asymptotically for any delay. $\square$

Remark 7.3. *In Theorem 7.3 we have shown that by introducing a time delay τ in the term $\sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t)$, the consensus is still guaranteed. The protocol can be implemented in practice since no instantaneous control input information of neighbors is needed. However, the performance will be deteriorated for a large τ . Next, we will show that the consensus speed is continuously dependent on τ , which means that the difference between the consensus speed under protocol (7.22) and the one under nominal protocol (7.6) can be limited when τ is small enough.*

Lemma 7.4. *For the eigenvalues of matrix $(D + B)^{-1}A$, we have*

$$\mu_1 < 0, \mu_n > 0,$$

if and only if $A \neq 0$.

Proof. It is obvious that $(D + B)^{-1}A$ is a nonnegative matrix¹. Note that $(D + B)^{-1}A$ may not be irreducible. By recalling Perron-Frobenius theorem of reducible matrix [6], it is known that $\mu_n \geq 0$. We first assume that $A \neq 0$. If A is irreducible, then $\mu_n > 0$. When A is reducible, according to [103], there exists a permutation matrix P_1 such that

$$P_1(D + B)^{-1}AP'_1 = \begin{bmatrix} N_1 & & & \\ & \ddots & & \\ & & N_l & \\ & & & 0 \end{bmatrix},$$

where $N_i, i = 1, \dots, l$ are irreducible matrices. We can get this block diagonal form since $(D + B)^{-1}A$ has a symmetric structure, i.e. $[(D + B)^{-1}A]^{i,j} > 0$ if and only if $[(D + B)^{-1}A]^{j,i} > 0$. Hence $\{\mu_1, \dots, \mu_n\}$ consists of all the eigenvalues of $N_i, i = 1, \dots, l$ as well as 0. Since $N_i, i = 1, \dots, l$ are irreducible, the Perron-Frobenius eigenvalues, denoted by $\lambda_{PF,i}, i = 1, \dots, l$, must be positive. Therefore,

$$\mu_n = \max_i \lambda_{PF,i} > 0.$$

It is straightforward that $A = 0$ implies that $\mu_n = 0$. On the other hand, due to that $(D + B)^{-1}A$ is the normalized adjacency matrix associated with graph $\mathcal{G}$ (a subgraph of $\bar{\mathcal{G}}$), $\text{trace}[(D + B)^{-1}A] = 0$, which implies that $\mu_1 < 0$ when $\mu_n > 0$. The proof is completed. $\square$

¹A matrix is called nonnegative matrix if all of its elements are nonnegative numbers.

By introducing the rate of convergence [102]

$$R \triangleq \lim_{t \rightarrow \infty} \frac{-\ln \|x(t) - x^o\|_2}{t}, \quad (7.27)$$

we have the following result.

Theorem 7.4. *Consider MAS (7.1) with undirected communication graph $\mathcal{G}$. If the whole system can reach consensus under the protocol (7.22), then when $\tau \rightarrow 0^+$,*

$$R = \gamma + \frac{\gamma^2 \mu_1}{1 - \mu_1} \tau + O(\tau^2). \quad (7.28)$$

Proof. Taking Laplace transform of (7.24) results in the following equation

$$[sI - se^{-\tau s}(D + B)^{-1}A + \gamma(D + B)^{-1}(L + B)] \tilde{X}(s) = 0,$$

where $\tilde{X}(s)$ is the Laplace transform of $\tilde{x}(t)$. For the sake of analyzing the stability, we ignore the initial condition here. The characteristic equation of system (7.24) is

$$\det [sI - se^{-\tau s}(D + B)^{-1}A + \gamma(D + B)^{-1}(L + B)] = 0. \quad (7.29)$$

It is well known that the convergence rate of $\tilde{x}(t)$ is dominated by the supremum of the real component of the solution to (7.29), which can be simplified as follows

$$\begin{aligned} & \det [sI - se^{-\tau s}(D + B)^{-1}A + \gamma(D + B)^{-1}(L + B)] \\ &= \det [(D + B)^{-1} (s(D + B) - se^{-\tau s}A + \gamma(L + B))] \\ &= \det [(D + B)^{-1} ((s + \gamma)(D + B) - (se^{-\tau s} + \gamma)A)] \\ &= \det [(s + \gamma)I - (se^{-\tau s} + \gamma)(D + B)^{-1}A] \\ &= \prod_{i=1}^n [(s + \gamma) - \mu_i(se^{-\tau s} + \gamma)]. \end{aligned} \quad (7.30)$$

In view of (7.25), note that when $\tau = 0$, there is only one solution $s = -\gamma$. Now we consider $\tau \neq 0$. There are n terms in the product in equation (7.30). We analyze

one of them first. Let $s_i = f + \iota g$, $f, g \in \mathbb{R}$ be the solution to

$$(s + \gamma) - \mu_i(se^{-\tau s} + \gamma) = 0,$$

then we have

$$(f + \iota g + \gamma) - \mu_i [e^{-\tau f}(f \cos \tau g + g \sin \tau g) + \iota e^{-\tau f}(g \cos \tau g - f \sin \tau g) + \gamma] = 0,$$

which results in the following two equations

$$\begin{aligned} f + \gamma(1 - \mu_i) - \mu_i e^{-\tau f}(f \cos \tau g + g \sin \tau g) &= 0, \\ g - \mu_i e^{-\tau f}(g \cos \tau g - f \sin \tau g) &= 0. \end{aligned}$$

Now let's consider two functions

$$\begin{aligned} H_1(f, g, \tau) &= f + \gamma(1 - \mu_i) - \mu_i e^{-\tau f}(f \cos \tau g + g \sin \tau g), \\ H_2(f, g, \tau) &= g - \mu_i e^{-\tau f}(g \cos \tau g - f \sin \tau g). \end{aligned}$$

Suppose $(f, g, \tau) \in \mathbb{D} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, where $\mathbb{D}$ is the domain of definition of (f, g, τ) . It is easy to check that the first order partial derivatives of H_1 and H_2 with respect to f, g and τ are continuous. Then we consider the following Jacobian determinant at point $(f, g, \tau) = (-\gamma, 0, 0)$,

$$\frac{\partial(H_1, H_2)}{\partial(f, g)} \Big|_{(-\gamma, 0, 0)} \triangleq \begin{vmatrix} \frac{\partial H_1}{\partial f} & \frac{\partial H_1}{\partial g} \\ \frac{\partial H_2}{\partial f} & \frac{\partial H_2}{\partial g} \end{vmatrix} \Big|_{(-\gamma, 0, 0)} = (1 - \mu_i)^2 \neq 0.$$

Since the Jacobian determinant is non-zero, there exists a neighborhood U_1 at point $(f, g) = (-\gamma, 0)$ and a neighborhood U_2 at point $\tau = 0$ such that in the neighborhood $U_1 \times U_2$ of point $(-\gamma, 0, 0) \in \mathbb{D}$, $H_1(f, g, \tau) = 0$ and $H_2(f, g, \tau) = 0$ can uniquely determine two first order differentiable functions over U_2 :

$$f = h_1(\tau), \quad g = h_2(\tau).$$

By simple calculation, we can see that the second order derivatives of h_1 and h_2 with respect to τ are bounded. Next, we calculate the gradient with respect to τ :

$$\dot{h}_1(0) = -\frac{\partial H_1}{\partial \tau} \bigg/ \frac{\partial H_1}{\partial f} \bigg|_{\tau=0} = -\frac{\gamma^2 \mu_i}{1 - \mu_i}, \quad \dot{h}_2(0) = -\frac{\partial H_1}{\partial \tau} \bigg/ \frac{\partial H_1}{\partial g} \bigg|_{\tau=0} = 0.$$

Due to the boundedness of the second order derivatives of h_1 and h_2 , we have the following Taylor expansion,

$$\begin{aligned} h_1(\tau) &= h_1(0) + \dot{h}_1(0)\tau + O(\tau^2) = -\gamma - \frac{\gamma^2 \mu_i}{1 - \mu_i} \tau + O(\tau^2), \\ h_2(\tau) &= h_2(0) + \dot{h}_2(0)\tau + O(\tau^2) = O(\tau^2). \end{aligned}$$

Hence

$$\operatorname{Re}(s_i) = -\gamma - \frac{\gamma^2 \mu_i}{1 - \mu_i} \tau + O(\tau^2).$$

For any $\mu_i > \mu_j$, we have $\operatorname{Re}(s_i) < \operatorname{Re}(s_j)$ if τ is small enough. Therefore, for

$$\sigma = \{s \in \mathbb{C} \mid s \text{ satisfies (7.29)}\}, \quad (7.31)$$

we have

$$\sup_{s \in \sigma} \operatorname{Re}(s) = -\gamma - \frac{\gamma^2 \mu_1}{1 - \mu_1} \tau + O(\tau^2). \quad (7.32)$$

Define the Hilbert space

$$\mathcal{H} = \mathbb{R}^n \times L_2\{-\tau, 0; \mathbb{R}^n\}, \quad (7.33)$$

with the obvious inner product. We introduce a semigroup $\mathcal{S}(\cdot)$ over space $\mathcal{H}$ based on system (7.24), which is defined as

$$\mathcal{X}(t) = \mathcal{S}(t)\mathcal{X}(0), \quad (7.34)$$

where

$$\mathcal{X}(t) \triangleq (\tilde{x}(t), \tilde{x}(t + \theta)), \quad \theta \in [-\tau, 0]. \quad (7.35)$$

It can be checked that $\mathcal{S}$ is differentiable and the corresponding infinitesimal gener-

ator $\mathcal{A}$ is given as

$$\mathcal{A}(w, z) = ((D + B)^{-1} A \dot{z}_\theta(\tau) - \gamma(D + B)^{-1}(L + B)w, z_\theta),$$

where $\dot{z}_\theta(\tau) = \frac{dz}{d\theta}|_{\theta=\tau}$. By tedious calculation, we can obtain the characteristic function of $\mathcal{A}$ which is consistent with (7.29). By recalling [8], the type of $\mathcal{S}$ is given as follows

$$\omega(\mathcal{S}) = \inf_{t>0} \frac{\ln \|\mathcal{S}(t)\|_2}{t} = \lim_{t \rightarrow \infty} \frac{\ln \|\mathcal{S}(t)\|_2}{t}. \quad (7.36)$$

Since $\mathcal{S}$ is differentiable, we have

$$\omega(\mathcal{S}) = \sup_{s \in \sigma} \operatorname{Re}(s). \quad (7.37)$$

By comparing (7.27), (7.36) and considering (7.32) and (7.37), (7.28) is obtained. $\square$

Remark 7.4. According to Lemma 7.4 and Theorem 7.4, $R > -\gamma$ for small τ , i.e. the consensus slows down after introducing delay. For the trivial case $A = 0$, which means every agent is not connected to any other agents except the reference state, the protocol (7.22) becomes

$$u_i(t) = -\gamma(x_i(t) - x^o),$$

which does not contain neighbors' information. Therefore, delay does not affect the consensus speed in this case.

Next, we consider the case that $\dot{x}^o(t) = v(t)$, which means that all the agents travel with a desired speeding profile $v(t)$ while synchronized. We assume that $v(t)$ is a Lipschitz continuous function with Lipschitz constant C . A modified protocol is proposed

$$u_i(t) = \frac{1}{d_i + b_i} \left\{ b_i v(t) - \gamma \left[\sum_{j \in \mathcal{V}} a_{i,j} (x_i(t) - x_j(t)) + b_i (x_i(t) - x^o(t)) \right] + \sum_{j \in \mathcal{V}} a_{i,j} \dot{x}_j(t - \tau) \right\}. \quad (7.38)$$

Theorem 7.5. *Consider MAS (7.1) with undirected communication graph $\mathcal{G}$. If the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree and $\dot{x}^o(t) = v(t)$, then under the protocol (7.38)*

$$\max_i \limsup_{t \rightarrow \infty} |x_i(t) - x^o(t)| \leq \delta_1(C, \tau, \gamma), \quad (7.39)$$

$$\max_i \limsup_{t \rightarrow \infty} |u_i(t) - v(t)| \leq \delta_2(C, \tau, \gamma), \quad (7.40)$$

where δ_1 and δ_2 are finite-value functions which satisfy

$$\lim_{C \rightarrow 0} \delta_1(C, \tau, \gamma) = 0, \quad \lim_{\tau \rightarrow 0} \delta_1(C, \tau, \gamma) = 0, \quad (7.41)$$

$$\lim_{C \rightarrow 0} \delta_2(C, \tau, \gamma) = 0, \quad \lim_{\tau \rightarrow 0} \delta_2(C, \tau, \gamma) = 0. \quad (7.42)$$

Proof. We rewrite the whole system as follows

$$\begin{cases} \dot{x}(t) = u(t), \\ u(t) = (D + B)^{-1} \{ Au(t - \tau) - \gamma [(L + B)x(t) - bx^o(t)] + bv(t) \}. \end{cases}$$

By considering the first equation and $\tilde{x}(t)$ defined as in (7.7), the second equation can be written as

$$\dot{\tilde{x}}(t) = (D + B)^{-1} A \dot{\tilde{x}}(t - \tau) - \gamma (D + B)^{-1} (L + B) \tilde{x}(t) + A \mathbf{1}_n (v(t - \tau) - v(t)). \quad (7.43)$$

Let $\xi(t) = (A \mathbf{1}_n (v(t - \tau) - v(t)), 0) \in \mathcal{H}$, then

$$\mathcal{X}(t) = \mathcal{S}(t) \mathcal{X}(0) + \int_0^t \mathcal{S}(t - \theta) \xi(\theta) d\theta,$$

where $\mathcal{H}$, $\mathcal{X}$ and $\mathcal{S}$ are as defined in (7.33), (7.35) and (7.34), respectively. Taking the norm of $\mathcal{X}(t)$ we have

$$\|\mathcal{X}(t)\|_2 \leq \|\mathcal{S}(t)\|_2 \|\mathcal{X}(0)\|_2 + \int_0^t \|\mathcal{S}(t - \theta)\|_2 d\theta \sup_{0 \leq s \leq t} \|\xi(s)\|_2. \quad (7.44)$$

According to Theorem 7.3 and by considering (7.36), we have

$$\|\mathcal{S}(t)\|_2 \leq K_1 e^{\omega(\mathcal{S})t}, \quad (7.45)$$

where K_1 is a positive number. By (7.37), we know that $\omega(\mathcal{S}) < 0$ for all $\tau \geq 0$.

Therefore,

$$\lim_{t \rightarrow \infty} \int_0^t \|\mathcal{S}(\theta)\|_2 d\theta = \frac{-K_1}{\omega(\mathcal{S})}. \quad (7.46)$$

Since $v(t)$ is Lipschitz continuous, it is easy to check that

$$\sup_{s \geq 0} \|\xi(s)\|_2 \leq K_2 C \tau, \quad (7.47)$$

where K_2 is a positive number. Then by considering (7.44)-(7.47) we have

$$\limsup_{t \rightarrow \infty} \|\mathcal{X}(t)\|_2 \leq -\frac{K_1 K_2 C \tau}{\omega(\mathcal{S})}.$$

By the definition of $\mathcal{X}(t)$ in (7.35), we know that

$$\|\mathcal{X}(t)\|_2^2 = \|\tilde{x}(t)\|_2^2 + \int_{-\tau}^0 \|\tilde{x}(t + \theta)\|_2^2 d\theta.$$

Moreover,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathcal{X}(t)\|_2^2 &= \limsup_{t \rightarrow \infty} \left(\|\tilde{x}(t)\|_2^2 + \int_{-\tau}^0 \|\tilde{x}(t + \theta)\|_2^2 d\theta \right) \\ &= (\tau + 1) \limsup_{t \rightarrow \infty} \|\tilde{x}(t)\|_2^2, \end{aligned}$$

which results in that

$$\limsup_{t \rightarrow \infty} \|\tilde{x}(t)\|_2 \leq -\frac{K_1 K_2 C \tau}{\sqrt{\tau + 1} \omega(\mathcal{S})} \triangleq \delta_1(C, \tau, \gamma). \quad (7.48)$$

δ_1 is a function of γ because $\omega(\mathcal{S})$ is only determined by γ and τ once the network

is fixed. On the other hand,

$$\max_i \limsup_{t \rightarrow \infty} |x_i(t) - x^o(t)| \leq \limsup_{t \rightarrow \infty} \|\tilde{x}(t)\|_2,$$

which leads to (7.39). (7.41) is straightforward by considering that $\lim_{\tau \rightarrow 0} (-\omega(\mathcal{S})) = \gamma > 0$.

What remains is to prove the uniform boundedness of deviation in velocity. Denote

$$\xi_1(t) = A\mathbf{1}_n(v(t - \tau) - v(t)) - \gamma(D + B)^{-1}(L + B)\tilde{x}(t).$$

Taking both (7.47) and (7.48) into consideration, there must exist $\delta(C, \tau, \gamma)$ such that

$$\sup_{t \geq 0} \|\xi_1(t)\|_2 \leq \delta(C, \tau, \gamma),$$

where $\delta(C, \tau, \gamma)$ satisfies that

$$\lim_{C \rightarrow 0} \delta(C, \tau, \gamma) = 0, \quad \lim_{\tau \rightarrow 0} \delta(C, \tau, \gamma) = 0.$$

Define $c(t) = \dot{x}(t)$, then according to (7.43) we have

$$c(t) = (D + B)^{-1}Ac(t - \tau) + \xi_1(t). \quad (7.49)$$

According to the proof of Theorem 7.3, we know that all the eigenvalues of $(D + B)^{-1}A$ strictly lie in the unit disk, which means that the discrete-time system (7.49) is stable. Note that for any $t > 0$, there exists $r \in [0, \tau)$ such that $t + r = k\tau$, where k is an integer. Then we have

$$c(t) = [(D + B)^{-1}A]^k c(-r) + \sum_{i=1}^k [(D + B)^{-1}A]^{k-i} \xi_1(i\tau - r).$$

It is known that

$$\lim_{k \rightarrow \infty} \left\| \sum_{i=1}^k [(D + B)^{-1}A]^{k-i} \right\|_2 \leq K_3,$$

where K_3 is a bounded number. Hence

$$\limsup_{t \rightarrow \infty} \|c(t)\|_2 \leq K_3 \delta(C, \tau, \gamma) \triangleq \delta_2(C, \tau, \gamma).$$

Since

$$\max_i \limsup_{t \rightarrow \infty} |u_i(t) - v(t)| \leq \limsup_{t \rightarrow \infty} \|c(t)\|_2,$$

(7.40) and (7.42) follow immediately. $\square$

Remark 7.5. *Theorem 7.5 provides a uniform bound for the consensus error both in position and velocity. In fact, the bound of the error depends on how the velocity of reference state changes. This is understandable since each agent may not directly access the reference state. This is in fact a second order consensus achieved by MASs with single integrator dynamics. In particular, when the velocity of reference state is a constant, i.e. $C = 0$, the error disappears, which means both position and velocity are synchronized.*

7.2.3 Application to Formation Control

Instead of driving all the states of the agents to the reference state, we try to make each state of agent have a desired deviation from the reference state. All the desired deviations may be time varying and different from each other. The technique of extension to formation control is widely used in many existing works, e.g. [63], [97], [129], etc. By applying the same manipulation, we shall provide an application of our obtained results to the formation control problem.

We use $\Delta_i(t)$ to denote the state deviation of agent i . Then the formation is achieved when

$$\lim_{t \rightarrow \infty} (x_i(t) - \Delta_i(t) - x^o) = 0, \quad \forall i = 1, \dots, n. \quad (7.50)$$

In what follows, the synchronization results are extended to the formation control problem.

Corollary 7.1. *Consider multi-agent system (7.1) with undirected communication graph $\mathcal{G}$. If the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree and $\dot{x}^o(t) = 0$, then (7.50) holds under the protocol*

$$u_i(t) = \frac{1}{d_i + b_i} \left\{ -\gamma \left[\sum_{j \in \mathcal{V}} a_{i,j} (x_i(t) - x_j(t) - \Delta_{i,j}(t)) + b_i (x_i(t) - \Delta_i(t) - x^o) \right] + \sum_{j \in \mathcal{V}} a_{i,j} \left(\dot{x}_j(t - \tau) - \dot{\Delta}_j(t - \tau) \right) \right\} + \dot{\Delta}_i(t), \quad (7.51)$$

where $\Delta_{i,j}(t) \triangleq \Delta_i(t) - \Delta_j(t)$.

Proof. Define

$$x_{\delta,i}(t) = x_i(t) - \Delta_i(t), \quad u_{\delta,i}(t) = u_i(t) - \dot{\Delta}_i(t). \quad (7.52)$$

Note that (7.23) still holds by replacing $x_i(t)$ and $u_i(t)$ with $x_{\delta,i}(t)$ and $u_{\delta,i}(t)$, respectively. According to the proof of Theorem 7.3, we know that

$$\lim_{t \rightarrow \infty} (x_{\delta,i}(t) - x^o) = 0,$$

which implies (7.50). □

Corollary 7.2. *Consider multi-agent system (7.1) with undirected communication graph $\mathcal{G}$. The reference state satisfies $\dot{x}^o(t) = v(t)$, where $v(t)$ is a Lipschitz continuous function with Lipschitz constant C . If the augmented graph $\bar{\mathcal{G}}$ contains a directed spanning tree, then under the following protocol*

$$u_i(t) = \frac{1}{d_i + b_i} \left\{ b_i v(t) - \gamma \left[\sum_{j \in \mathcal{V}} a_{i,j} (x_i(t) - x_j(t) - \Delta_{i,j}(t)) + b_i (x_i(t) - \Delta_i(t) - x^o) \right] + \sum_{j \in \mathcal{V}} a_{i,j} \left(\dot{x}_j(t - \tau) - \dot{\Delta}_j(t - \tau) \right) \right\} + \dot{\Delta}_i(t), \quad (7.53)$$

we have

$$\begin{aligned} \max_i \limsup_{t \rightarrow \infty} |x_i(t) - \Delta_i(t) - x^o(t)| &\leq \delta_1(C, \tau, \gamma), \\ \max_i \limsup_{t \rightarrow \infty} |u_i(t) - \dot{\Delta}_i(t) - v(t)| &\leq \delta_2(C, \tau, \gamma), \end{aligned}$$

where δ_1 and δ_2 are the same as in (7.39) and (7.40), respectively.

Proof. Note that (7.53) can be rewritten in the same form as (7.38) with $x_i(t)$ and $u_i(t)$ replaced by $x_{\delta,i}(t)$ and $u_{\delta,i}(t)$ of (7.52), respectively. According to Theorem 7.5, the result follows. $\square$

7.3 Simulations and Experiments

7.3.1 Consensus Problem

In this section, we shall give two examples to verify the proposed protocols (7.22) and (7.38). Consider two connected agents. As is shown in Fig. 7.1, only the first agent can get the reference information. To be specific,

$$\begin{aligned} \bar{A} &= \left[\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right], \bar{D} = \left[\begin{array}{c|c} 2 & \\ \hline 1 & \\ 0 & \end{array} \right], \bar{L} = \left[\begin{array}{cc|c} 2 & -1 & -1 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array} \right], \\ A &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, D = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & \\ & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, L = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \end{aligned}$$

The initial state values are $x_1(0) = 3$, $x_2(0) = -1$. The reference state is $x^o = 1$. The two protocols (7.5) and (7.6) will be applied. According to Theorem 7.1, if $\kappa_1 = \kappa_2 = 1$, the optimal gains can be calculated as $\beta^* = \sqrt{0.4}$, $\gamma^* = 1$. Fig. 7.2

Figure 7.1: Augmented Graph $\bar{\mathcal{G}}$

Figure 7.2: Dot line: state trajectories under protocol (7.6); solid line: state trajectories under protocol (7.5)

shows the state trajectories of the two agents under protocol (7.5) and (7.6). We also get

$$\int_0^{25} \sum_{i=1}^n (d_i + b_i) (\tilde{x}_i^2(t) + u_{1,i}^2(t)) dt \Big|_{\beta=\beta^*} \approx 12.649 = J_1,$$

$$\int_0^{25} \sum_{i=1}^n (d_i + b_i) (\tilde{x}_i^2(t) + u_{2,i}^2(t)) dt \Big|_{\gamma=\gamma^*} \approx 12 = J_2.$$

Here $J_1 > J_2$ is due to that the two eigenvalues of $(D + B)^{-1}(L + B)$, $\eta_1 = 0.2929$ and $\eta_2 = 1.7071$, are different.

Next, we shall verify the protocols with an introduced delay in (7.22), (7.38). γ

is still set to be 1. First, we consider a constant reference $x^o = 1$. By applying protocol (7.22), we compare the cases $\tau = 0.2s$ and $\tau = 2s$ with the nominal case $\tau = 0s$ in Fig. 7.3 and Fig. 7.4, respectively. We can see that when τ is small, the trajectories of the states do not deviate too much from the nominal ones. However, when τ is large, the synchronization slows down.

Figure 7.3: Dot line: $\tau = 0s$; solid line: $\tau = 0.2s$

Then, we let $x^o(t) = v^o t + 0.5$, and $v^o = 0.1$. By applying protocol (7.38), the two agents achieve synchronization when $\tau = 0s$, $\tau = 0.2s$ and $\tau = 2s$. This is because $v(t)$ is a constant, which means $C = 0$, and therefore $\delta_1 = \delta_2 = 0$. The state trajectories are shown in Fig. 7.5 and Fig. 7.6.

Now we consider

$$x^o(t) = 1 + \sin(0.5t), \quad v(t) = 0.5 \cos(0.5t).$$

Let $\tau = 0.2s$. By applying the same controller, the trajectories are shown in Fig. 7.7. From the figure we can see that there are some deviations between the states of the agents and the reference state. This is because only the delayed control input

Figure 7.4: Dot line: $\tau = 0s$; solid line: $\tau = 2s$

Figure 7.5: Dot line: $\tau = 0s$; solid line: $\tau = 0.2s$

signals rather than instantaneous control input signals from neighbors are available. When $\tau = 0$, the deviation will disappear.

Figure 7.6: Dot line: $\tau = 0s$; solid line: $\tau = 2s$

Figure 7.7: Dot line: Trajectories of the two agents; solid line: Trajectory of reference state

7.3.2 Formation Problem

At last, we consider the formation problem. The communication topology in Fig. 7.1 is still applied. We shall verify the protocols (7.51) and (7.53) by both MATLAB

simulation and experiment. The simulation is carried out in 2-dimensional space. Since x and y coordinates are decoupled, the algorithms can be applied separately. Denote $x_i(t)$, $y_i(t)$ and $\Delta_{i,x}$, $\Delta_{i,y}$ ($i = 1, 2$) the real states and formation information of the two agents in both x and y directions. The reference state becomes (x^o, y^o) . We set $\tau = 0.1s$. In order to avoid saturation in the experiment, γ is chosen to be 0.1. We first carry out the MATLAB simulation. Given the initial positions of the two agents

$$x_1(0) = 600\text{cm}, y_1(0) = 200\text{cm}, x_2(0) = 400\text{cm}, y_2(0) = 200\text{cm},$$

we adopt static reference state

$$x^o = 450\text{cm}, y^o = 300\text{cm}. \quad (7.54)$$

Let

$$\Delta_{1,x} = 50\text{cm}, \Delta_{1,y} = -50\text{cm}, \Delta_{2,x} = -50\text{cm}, \Delta_{2,y} = 50\text{cm}.$$

By applying protocol (7.51), the formation in 2-dimensional space is achieved. Fig. 7.8 and Fig. 7.9 show the simulation result.

Then we let

$$x^o(t) = (500 - 5\sqrt{2}t)\text{cm}, y^o(t) = (100 + 5\sqrt{2}t)\text{cm}. \quad (7.55)$$

In this case, we set

$$\Delta_{1,x} = 50\text{cm}, \Delta_{1,y} = 0\text{cm}, \Delta_{2,x} = 0\text{cm}, \Delta_{2,y} = -50\text{cm}.$$

By applying protocol (7.53), the formation is shown in Fig. 7.10.

Next, we shall implement the algorithms on the Amigo robots, the model of which is given in Fig. 7.11. The kinematic equations for the i -th robot are

$$\dot{x}_i(t) = v_i(t) \cos(\theta_i(t)), \dot{y}_i(t) = v_i(t) \sin(\theta_i(t)), \dot{\theta}_i(t) = w(t),$$

Figure 7.8: x coordinates of the two agents and reference state Figure 7.9: y coordinates of the two agents and reference state

where x_i and y_i are the coordinates of the center of the two wheels, θ_i is the orientation of the robot, v_i and w_i are linear and angular velocity of the robot. We consider the coordinates of the hand h_x , h_y rather than x_i , y_i since a linear model can be obtained by a well-known transformation [93]. To be specific, it is easy to

Figure 7.10: Trajectories of the two agents and the reference state

Figure 7.11: Model of Amigo robot

check that for the i -th robot

$$\begin{bmatrix} h_{x,i}(t) \\ h_{y,i}(t) \end{bmatrix} = \begin{bmatrix} x_i(t) \\ y_i(t) \end{bmatrix} + l \begin{bmatrix} \cos(\theta_i(t)) \\ \sin(\theta_i(t)) \end{bmatrix},$$

and

$$\begin{bmatrix} \dot{h}_{x,i}(t) \\ \dot{h}_{y,i}(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta_i(t)) & -l \sin(\theta_i(t)) \\ \sin(\theta_i(t)) & l \cos(\theta_i(t)) \end{bmatrix} \begin{bmatrix} v_i(t) \\ w_i(t) \end{bmatrix},$$

where $l = 6.5\text{cm}$ is the length of the hand. By defining the new control inputs

$$\begin{aligned} u_{x,i}(t) &= \cos(\theta_i(t)) v_i(t) - l \sin(\theta_i(t)) w_i(t), \\ u_{y,i}(t) &= \sin(\theta_i(t)) v_i(t) + l \cos(\theta_i(t)) w_i(t), \end{aligned}$$

the dynamics of the i -th robot becomes linear

$$\dot{h}_{x,i}(t) = u_{x,i}(t), \quad \dot{h}_{y,i}(t) = u_{y,i}(t).$$

The initial coordinates of the two robots are given as follows

$$h_{x,1}(0) = 598.76\text{cm}, \quad h_{y,1}(0) = 202.82\text{cm}, \quad h_{x,2}(0) = 400.66\text{cm}, \quad h_{y,2}(0) = 205.67\text{cm}.$$

Given the static reference state $x^o = 300\text{cm}$, $y^o = 300\text{cm}$ and dynamic reference state $x^o(t) = (500 - 5\sqrt{2}t)\text{cm}$, $y^o = (100 + 5\sqrt{2}t)\text{cm}$, Under the same configuration as in the simulation, the trajectories of the robots are shown in Fig. 7.12, Fig. 7.13 and Fig. 7.14. The fluctuations are due to the measurement noises.

Figure 7.12: x coordinates of the Amigos given static reference state (7.54)

Figure 7.13: y coordinates of the Amigos given static reference state (7.54)

Figure 7.14: Formation by Amigo given dynamic reference state (7.55)

7.4 Conclusion

We have considered the consensus problem in an undirected network. Two protocols, (7.5) and (7.6), have been compared. It has been shown that protocol (7.6)

outperforms protocol (7.5) in some sense. Since the control inputs of each pair of neighbors are coupled in protocol (7.6), the implementation problem has been investigated. By introducing delays to the term which depends on neighbors' control inputs, we have proved that the consensus still can be achieved. Moreover, the effect of delay on the consensus speed has been studied. After that, an application to formation control was discussed. Some examples and experiments have been given to verify the results. It is worthy mentioning that the proposed results are expected to be extended to MASs with dynamic topologies.

Chapter 8

Distributed Consensus for MAS with Delays and Noises in Transmission Channels

In this chapter, we consider the consensus problem for linear discrete-time MASs in the presence of both transmission delays and noises. To overcome the challenges induced by both the transmission delays and noises, an auxiliary system is introduced for the first time. Based on the auxiliary system, the problem is transferred into that of analyzing the consensus for a delay-free system with input of a color noise for which the algebraic theory can then be applied.

Under the proposed consensus protocol, necessary and sufficient conditions are respectively obtained for both non-leader-follower and leader-follower cases under a fixed topology, which guarantee both strong consensus and mean square consensus. The condition consists of the persistence condition on weights and a connectivity condition on graph which states that the quotient graph contains a spanning tree. For the non-leader-follower case, our necessary and sufficient consensus conditions are the same as those for delay-free systems [65].

For the leader-follower case, our result is stronger than those of existing works even for delay-free case; see [75]. For dynamically switching topologies, we assume

that the union graph contains a spanning tree which guarantees that the graph is connected in the long term. For randomly switching topologies, the counterpart condition is that the mean graph is connected. Sufficient conditions for strong consensus and mean square consensus under both the dynamically switching topologies and randomly switching topologies are given. Some numerical examples are given to demonstrate the effectiveness of the protocol.

8.1 Problem Statement

We shall formulate the consensus problem to be studied for MASs. The system to be considered consists of n autonomous agents, labeled from 1 to n . We represent the communication network by a graph $\mathcal{G}$, and the vertex set is therefore given as $\mathcal{V}(\mathcal{G}) = \{1, \dots, n\}$. We shall extend the concept of neighborhood to the delay system setting. We still use $\mathcal{N}_i$ to denote the set of the vertices whose information (including delayed information) is received by vertex i . The agent i is assumed to have the following dynamics

$$x_i(k+1) = x_i(k) + u_i(k), \quad i \in \mathcal{V}, \quad (8.1)$$

where $x_i \in \mathbb{R}$ and $u_i \in \mathbb{R}$ are the state and input of agent i , respectively. It should be noted that although only the scalar systems are considered in this chapter, our results to be developed later can be extended to the vector valued state case by introducing Kronecker product. In networked systems, the neighbors' information is often delayed due to transmission delay. Denote $d_{i,j}(k) \in \bar{\mathbb{Z}}^+$ the transmission delay from agent j to i at time instant k . We assume that the time delay is time-varying with an upper bound $\bar{d}$, i.e. $d_{i,j}(k) \leq \bar{d}$, $\forall i, j \in \mathcal{V}$ and $\forall k \in \bar{\mathbb{Z}}^+$. On the other hand, in most cases, we cannot get the exact information from neighbors due to, for example, quantization errors and channel noises. Hence, we denote the received state value of j by i as

$$y_{i,j}(k) = x_j(k - d_{i,j}(k)) + w_{i,j}(k), \quad j \in \mathcal{N}_i. \quad (8.2)$$

The noise processes $\{w_{i,j}(k), k \in \bar{\mathbb{Z}}^+\}$ represent unreliable information exchanges in the network which satisfy the following assumption:

Assumption 8.1. $\{w_{i,j}(k), k \in \bar{\mathbb{Z}}^+, i, j \in \mathcal{V}\}$ are random processes with zero means and uniformly bounded covariances and are independent of the initial conditions $x_i(0)$, $i \in \mathcal{V}$. Furthermore, they are also temporally mutually independent¹, i.e.

$$E w_{i,j}(k_1) w_{i',j'}(k_2) = 0, \quad \forall k_1 \neq k_2, \quad i, j, i', j' \in \mathcal{V}.$$

Our objective is to design $u_i(k) \in \mathcal{L}(\{x_i(k - \tau), y_{i,j}(k) : \tau = 0, \dots, \bar{d}, j \in \mathcal{N}_i\})$ such that the states of all the agents achieve an agreement in a stochastic sense when $k \rightarrow \infty$, where $\mathcal{L}(S)$ denotes the linear space spanned by S . We introduce τ to make the distributed protocol more flexible. Note that in some networked systems, there may be delays for an agent to acquire its own state information (called inner-agent transmission delays). There are different definitions to characterize the asymptotic behavior of the agents [115, 48]:

Definition 8.1. (*Strong consensus*): The agents are said to reach strong consensus if there exists a random variable x^* such that $\forall i \in \mathcal{V}$,

$$\lim_{k \rightarrow \infty} x_i(k) = x^* \text{ a.s.} \quad (8.3)$$

Definition 8.2. (*Mean square consensus*): The MAS is said to reach mean square consensus if there exists a random variable x^* such that

$$\max_{i \in \mathcal{V}} \sup_{k \in \mathbb{Z}^+} E |x_i(k)|^2 < +\infty, \quad \lim_{k \rightarrow \infty} E |x_i(k) - x^*|^2 = 0, \quad \forall i \in \mathcal{V}. \quad (8.4)$$

Remark 8.1. For the delay-free case, i.e. $d_{i,j}(k) = 0$, [46], [66] and [48] analyze the consensus problem with different methods. Due to the presence of delays, the monotonically decreasing property of the Lyapunov function in [46] is no longer the case

¹The mutual independence assumption may be relaxed. For instance, the noise sequences can be assumed as martingale difference sequences. The independence assumption is only imposed on time and the noises at the same time instant are allowed to be dependent on each other.

and thus the supermartingale convergence analysis in [66] cannot be applied. The stochastic double array analysis in [48] can only be applied for some special graphs such as networks with symmetric structures and without communication delays.

In the rest of this chapter, we shall develop a new approach for MAS (8.1)-(8.2) to achieve both the strong consensus and mean square consensus.

8.2 Distributed Consensus

8.2.1 Fixed Topology

We consider the MAS with time-varying communication delays (8.1)-(8.2). The control protocol for the i -th agent is proposed as

$$u_i(k) = c(k) \sum_{j \in \mathcal{V}} a_{i,j} (y_{i,j}(k) - x_i(k - t_{i,j}(k))), \quad (8.5)$$

where $a_{i,j} \in \mathcal{C} \cup \{0\}$ with $\mathcal{C}$ a compact set on $\mathbb{R}^+$, $c(k) \in \bar{\mathbb{R}}^+$ is the consensus gain function [66] and $t_{i,j}(k) \in \bar{\mathbb{Z}}^+$ is the inner-agent delay satisfying $0 \leq t_{i,j}(k) \leq \bar{d}$. We shall prove that the proposed protocol is robust with respect to both the inter-agent transmission delays $d_{i,j}(k)$ and inner-agent delays $t_{i,j}(k)$. Define

$$X(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix}, \quad W(k) = \Xi \begin{bmatrix} \bar{w}_1(k) \\ \vdots \\ \bar{w}_n(k) \end{bmatrix} \triangleq \begin{bmatrix} w_1(k) \\ \vdots \\ w_n(k) \end{bmatrix}, \quad (8.6)$$

where

$$\Xi = \text{diag}\{\bar{a}_1, \dots, \bar{a}_n\}, \quad \bar{a}_i = [a_{i,1} \ \cdots \ a_{i,n}], \quad \bar{w}_i(k) = \text{col}\{w_{i,1}(k), \dots, w_{i,n}(k)\}. \quad (8.7)$$

It is obvious that $w_i(k) = \sum_{j \in \mathcal{N}_i} a_{i,j} w_{i,j}(k)$, $i \in \mathcal{V}$. Then the closed-loop system can be written in a compact form as

$$X(k+1) = X(k) - c(k) \sum_{l=0}^{\bar{d}} L_l(k) X(k-l) + c(k) W(k), \quad (8.8)$$

where

$$L_l^{i,j}(k) \triangleq \begin{cases} -a_{i,j} \delta_{d_{i,j}(k),l}, & i \neq j, \\ \sum_{m \in \mathcal{V} \setminus \{i\}} a_{i,m} \delta_{t_{i,m}(k),l}, & i = j. \end{cases}$$

Due to the multiple state delays and time-varying matrix $L_l(k)$ in system (8.8), the stability is difficult to analyze. To overcome the difficulty, we introduce the following auxiliary system,

$$\begin{cases} X_1(k+1) = A(k)X_1(k) + W_c(k), \\ W_c(k) = \sum_{i=1}^{\bar{d}} B_i(k)W_c(k-i) + c(k)W(k), \end{cases} \quad (8.9)$$

where $A(k)$ and $B_i(k)$, $i = 1, \dots, \bar{d}$ satisfy

$$A(k) + B_1(k) = I - c(k)L_0(k), \quad (8.10)$$

$$B_1(k)A(k-1) - B_2(k) = c(k)L_1(k), \quad (8.11)$$

$\vdots$

$$B_{\bar{d}-1}(k)A(k-\bar{d}+1) - B_{\bar{d}}(k) = c(k)L_{\bar{d}-1}(k), \quad (8.12)$$

$$B_{\bar{d}}(k)A(k-\bar{d}) = c(k)L_{\bar{d}}(k). \quad (8.13)$$

By comparing the auxiliary system (8.9) with (8.8), it is clear that $X(k) = X_1(k)$ if their initial conditions are the same. Further, observe that the first system in (8.9) is delay-free with colored noise input $W_c(k)$.

Note that $\sum_{l=0}^{\bar{d}} L_l(k) = L$ is the Laplacian matrix of the quotient graph. Denote

$$A(k) = I - c(k)L + \Delta_A(k), \quad B_i(k) = \sum_{l=i}^{\bar{d}} c(k)L_l(k) + \Delta_{B,i}(k), \quad \rho_m(k) = \max_{0 \leq l \leq m} c^2(k-l). \quad (8.14)$$

The following lemma presents some properties of $A(k)$ and $B_i(k)$, $i = 1, \dots, \bar{d}$.

Lemma 8.1. *Consider system (8.9). If $c(k)$ satisfies $\lim_{k \rightarrow \infty} c(k) = 0$, there exist time instant k_I and proper initial conditions on $A(k_I - l)$, $l = 1, \dots, \bar{d}$ such that $\forall k \geq k_I$, $A(k)$ is invertible. Moreover, there exists $C > 0$ such that $\forall k \geq k_I$*

$$\|\Delta_A(k)\|_2 \leq C\rho_{\bar{d}}(k), \quad \|\Delta_{B,i}(k)\|_2 \leq C\rho_{\bar{d}}(k), \quad (8.15)$$

where $\Delta_A(k)$, $\Delta_{B,i}(k)$ and $\rho_{\bar{d}}(k)$ are as defined in (8.14).

The following lemma is needed to prove Lemma 8.1.

Lemma 8.2. ([142]) *For any matrix $M = I - \Delta$ satisfying $\|\Delta\|_2 \leq \alpha$ with $0 < \alpha < 1$, we have that M is invertible and $\|M^{-1}\|_2 \leq (1 - \|\Delta\|_2)^{-1} \leq (1 - \alpha)^{-1}$.*

Proof of Lemma 8.1. Since $\lim_{k \rightarrow \infty} c(k) = 0$, there must exist $s : \mathbb{R}^+ \mapsto \mathbb{Z}^+$ such that $\forall K \in \mathbb{R}^+$,

$$\sup_{k \geq s(K)} c(k) \leq K. \quad (8.16)$$

Define

$$k_0 = s(K), \quad K = \frac{\sqrt{n}\|L\|_\infty}{C(C-1)} \wedge \frac{\alpha}{2\sqrt{n}\|L\|_\infty} \wedge \sqrt{\frac{\alpha}{2C}}, \quad C = 1 + n\|L\|_\infty^2 \sum_{l=0}^{\bar{d}} \frac{(1-\alpha)^{-l} - 1}{\alpha}, \quad (8.17)$$

where α is a constant satisfying $0 < \alpha < 1$. Then we choose $k_I = k_0 + \bar{d}$ and the initial values of $A(k)$ with $\Delta_A(k)$ being any matrix satisfying that $\forall k = k_I - 1, \dots, k_I - \bar{d}$,

$$\|\Delta_A(k)\|_2 \leq C\rho_{\bar{d}}(k). \quad (8.18)$$

It is easy to find such Δ_A , for example, we can choose $\Delta_A(k) = C\rho_{\bar{d}}(k)I$. In the light of (8.17)-(8.18) and $\|L\|_2 \leq \sqrt{n}\|L\|_\infty$, we have that $\forall k_I - \bar{d} \leq k \leq k_I - 1$,

$$\|c(k)L - \Delta_A(k)\|_2 \leq c(k)\|L\|_2 + \|\Delta_A(k)\|_2 \leq K\sqrt{n}\|L\|_\infty + CK^2 \leq \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha < 1. \quad (8.19)$$

Thus, it follows from Lemma 8.2 that $A(k_I - l)$, $1 \leq l \leq \bar{d}$ are invertible. Next, by considering (8.10)-(8.13), we have that

$$I - A(k_I) = \sum_{l=0}^{\bar{d}} c(k_I) L_l(k_I) (\Pi_{k_I-1, k_I-l}^A)^{-1} = c(k_I) L(k_I) - \Delta_A(k_I). \quad (8.20)$$

Denote $g(k) = \max_{l=1, \dots, \bar{d}} c(k-l) \sqrt{n} \|L\|_\infty + C \rho_{\bar{d}}(k)$, then $\forall l = 1, \dots, \bar{d}$,

$$\|c(k_I - l) L(k_I - l) - \Delta_A(k_I - l)\|_2 \leq g(k_I) \leq \alpha. \quad (8.21)$$

According to Lemma 8.2 and (8.20)-(8.21), we know that for all $l = 1, \dots, \bar{d}$,

$$\|A^{-1}(k_I - l)\|_2 \leq \frac{1}{1 - \|c(k_I - l) L(k_I - l) - \Delta_A(k_I - l)\|_2} \leq \frac{1}{1 - g(k_I)}.$$

Furthermore, it is easy to check that $\forall l \in \mathbb{Z}^+$,

$$\frac{(1 - g(k_I))^{-l} - 1}{g(k_I)} \leq \frac{(1 - \alpha)^{-l} - 1}{\alpha}. \quad (8.22)$$

From (8.20) we have the following inequality

$$\begin{aligned} \|\Delta_A(k_I)\|_2 &= \left\| \sum_{l=0}^{\bar{d}} c(k_I) L_l(k_I) \left[(\Pi_{k_I-1, k_I-l}^A)^{-1} - I \right] \right\|_2 \\ &\leq \sum_{l=0}^{\bar{d}} c(k_I) \|L_l(k_I)\|_2 \left[(1 - g(k_I))^{-l} - 1 \right] \\ &\leq \sum_{l=0}^{\bar{d}} c(k_I) g(k_I) \|L_l(k_I)\|_2 \frac{(1 - \alpha)^{-l} - 1}{\alpha} \\ &\leq \sum_{l=0}^{\bar{d}} (\rho_{\bar{d}}(k_I) \sqrt{n} \|L\|_\infty + KC \rho_{\bar{d}}(k_I)) \sqrt{n} \|L\|_\infty \frac{(1 - \alpha)^{-l} - 1}{\alpha} \\ &\leq \rho_{\bar{d}}(k_I) \left(C - 1 + CK \frac{C - 1}{\sqrt{n} \|L\|_\infty} \right) \\ &\leq C \rho_{\bar{d}}(k_I). \end{aligned} \quad (8.23)$$

In the above, the second inequality is due to (8.22), the third inequality is due to

the definition of $g(k)$ and the fact that $\|L_l(k)\|_2 \leq \sqrt{n}\|L_l(k)\|_\infty \leq \sqrt{n}\|L\|_\infty$, the last two inequalities follow from (8.17). Thus, it follows similarly that (8.19) holds for $k = k_I$ and $A(k_I)$ is invertible. By induction, we can arrive at that $\forall k > k_I$, $A(k)$ is invertible and $\|\Delta_A(k)\|_2 \leq C\rho_{\bar{d}}(k)$.

Next, by some simple calculation we have

$$B_i(k) = \sum_{l=i}^{\bar{d}} c(k)L_l(k) \left(\Pi_{k-i,k-l}^A\right)^{-1}, \quad \forall i = 1, 2, \dots, \bar{d}. \quad (8.24)$$

By comparing (8.14) and (8.24), it follows

$$\|\Delta_{B,i}(k)\|_2 = \left\| \sum_{l=i}^{\bar{d}} c(k)L_l(k) \left[\left(\Pi_{k-i,k-l}^A\right)^{-1} - I \right] \right\|_2 \leq C\rho_{\bar{d}}(k).$$

The inequality follows from the same manipulation as that in (8.23). $\square$

Since the auxiliary system (8.9) and the original system (8.8) are equivalent under the same initial conditions, we can solve the consensus problem by analyzing the system (8.9) instead. In the following, we shall give necessary and sufficient conditions under which the protocol (8.5) can make the MAS (8.1) reach strong consensus and mean square consensus. The following lemmas are essential to establishing our main results.

Lemma 8.3. *Given an integer $m > 0$, for any matrix sequence $\{M(k), k \geq m\}$ satisfying*

$$\|M(k) - I + c(k)L(k)\|_2 \leq C_1\rho_m(k),$$

where $C_1 > 0$ is a constant, $\rho_m(k)$ is as defined in (8.14), the following statements hold:

1) For $c(k)$ satisfying that

$$\sum_{k=0}^{\infty} c(k) = +\infty, \quad \lim_{k \rightarrow \infty} c(k) = 0, \quad (8.25)$$

if $L(k) = L$ is a constant matrix and the real component of any eigenvalue of L is positive, there exist $k_I \in \mathbb{Z}^+$, C_2 , C_3 , $C_4 > 0$ such that

$$\|\Pi_{k_2, k_1}^M\|_2 \leq C_2 \prod_{k=k_1}^{k_2} (1 - C_3 c(k) + C_4 \rho_m(k)), \quad \forall k_1 \geq k_I. \quad (8.26)$$

Furthermore, $\forall k_1 \geq m$, $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^M = 0$.

2) For $c(k)$ satisfying that

$$\sum_{k=0}^{\infty} c(k) = +\infty, \quad \sum_{k=0}^{\infty} c^2(k) < +\infty, \quad (8.27)$$

if $L(k) = \text{diag}\{0, L_{(1)}\}$ is a constant matrix and the real component of any eigenvalue of $L_{(1)}$ is positive, then for any $k_1 \geq m$, Π_{k_2, k_1}^M converges to a matrix with the last $n - 1$ rows equal to 0 as $k_2 \rightarrow \infty$. The convergence, however, does not hold when $\sum_{k=0}^{\infty} c(k) < +\infty$.

3) For $c(k)$ satisfying (8.27), if $L(k) = \text{diag}\{0, L_{(1)}(k)\}$ and the eigenvalues of $L_{(1)}(k) + L_{(1)}^T(k)$ are uniformly positive, then for any $k_1 \geq m$, Π_{k_2, k_1}^M converges to a matrix with the last $n - 1$ rows equal to 0 as $k_2 \rightarrow \infty$.

Proof. The proof can be found in the Appendix. $\square$

Lemma 8.4. For any $k_I \in \bar{\mathbb{Z}}^+$ and sequences $\{\gamma_1(k), k \geq k_I\}$, $\{\gamma_2(k), k \geq k_I\}$ satisfying $0 \leq \gamma_1(k) < 1$ and $\sum_{k=k_I}^{\infty} \gamma_1(k) = +\infty$, we have

$$\lim_{k \rightarrow \infty} \sum_{k_1=k_I}^k \gamma_2(k_1) \prod_{l=k_1+1}^k (1 - \gamma_1(l)) = \lim_{k \rightarrow \infty} \frac{\gamma_2(k)}{\gamma_1(k)}. \quad (8.28)$$

Proof. Consider system

$$\theta(k+1) = (1 - \gamma_1(k))\theta(k) + \gamma_2(k),$$

with initial condition $\theta(k_I) = 0$. It is clear that $\theta(k+1) = \sum_{k_1=k_I}^k \gamma_2(k_1) \prod_{l=k_1+1}^k (1 -$

$\gamma_1(l)$). On the other hand, by recalling Theorem 1.2.22 [40], there is $\lim_{k \rightarrow \infty} \theta(k) = \lim_{k \rightarrow \infty} \frac{\gamma_2(k)}{\gamma_1(k)}$, which implies (8.28). $\square$

Now, we are ready to give the main result. Since a topology with leader-follower structure is special, we shall discuss the consensus problem for the leader-follower case in separate theorems.

Theorem 8.1. (Non-leader-follower case) *Consider the MAS (8.1)-(8.2) with interaction protocol (8.5). For any noises satisfying Assumption 8.1 and any topology satisfying $\deg_{in}(i) > 0, \forall i \in \mathcal{V}$, the strong consensus is achieved if and only if the quotient graph contains a spanning tree and the consensus gain function satisfies (8.27).*

Proof. ‘Sufficiency’. Consider system (8.8). The MAS (8.1) reaches strong consensus is equivalent to that $\lim_{k \rightarrow \infty} X(k) = \mathbf{1}_n x^*$ almost surely, where x^* is as defined in (8.3). It is straightforward that $\sum_{i=0}^{\bar{d}} L_i(k) = L$, where L is the Laplacian matrix corresponding to the quotient graph. Denote $\{\lambda_i, i = 1, \dots, n\}$ the eigenvalues of L . We assume that $\text{Re}(\lambda_i) \leq \text{Re}(\lambda_j)$ when $i < j$. According to Lemma A.1, that the quotient graph contains a spanning tree implies that $\lambda_1 = 0$, $\text{Re}(\lambda_2) > 0$. So, there exists an invertible matrix T with the first column equal to $\mathbf{1}_n$ and $L = T \text{diag}\{0, L_{(1)}\} T^{-1}$, where $L_{(1)}$ is an $(n-1) \times (n-1)$ matrix with the eigenvalues $\lambda_2, \dots, \lambda_n$.

It is very difficult to analyze the convergence property of system (8.8) due to multiple delays in the state. However, according to Lemma 8.1, we can write system (8.8) in the form of (8.9) with initial time $k_I = k_0 + \bar{d}$ and initial values of A satisfying (8.18), where k_0 is as defined in (8.17). Next, we shall show that $\forall k_1 \geq k_I$, Π_{k_2, k_1}^A converges to a matrix with identical rows as $k_2 \rightarrow \infty$. To this end, in view of (8.14), $A(k)$ can be written as $A(k) = T A_1(k) T^{-1}$, where

$$A_1(k) = I - c(k) \text{diag}\{0, L_{(1)}\} + \Delta_{A_1}(k), \quad (8.29)$$

and $\Delta_{A_1}(k) = T^{-1} \Delta_A(k) T$. According to Lemma 8.1, $\|\Delta_{A_1}(k)\|_2 \leq C_T C \rho_{\bar{d}}(k)$, where $C_T = \|T\|_2 \|T^{-1}\|_2$, C is as defined in (8.17). Then by applying part 2) of

Lemma 8.3, $\Pi_{k_2, k_1}^{A_1}$ converges to a matrix with the last $n - 1$ rows equal to 0 as $k_2 \rightarrow \infty$. Hence,

$$\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^A = T \lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^{A_1} T^{-1} \triangleq T \begin{bmatrix} r_{1,1}(k_1) & r_{1,2}(k_1) & \cdots & r_{1,n}(k_1) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \mathbf{1}_n r_1(k_1), \quad (8.30)$$

where $r_1(k_1) = [r_{1,1}(k_1) \ r_{1,2}(k_1) \ \cdots \ r_{1,n}(k_1)]$.

Next, we write (8.9) into the following augmented system

$$z(k+1) = \Phi(k)z(k) + \Gamma c(k+1)W(k+1), \quad (8.31)$$

where $z(k) = [X'_1(k) \ W'_c(k) \ W'_c(k-1) \ \cdots \ W'_c(k-\bar{d}+1)]'$,

$$\Phi(k) = \begin{bmatrix} A(k) & D \\ 0 & B(k) \end{bmatrix}, \quad B(k) = \begin{bmatrix} B_1(k+1) & B_2(k+1) & \cdots & B_{\bar{d}}(k+1) \\ I_n & & & 0 \\ & \ddots & & \vdots \\ & & I_n & 0 \end{bmatrix}, \quad (8.32)$$

$$D = [I_n \ 0 \ 0 \ \cdots \ 0], \quad \Gamma = [0 \ D]'$$

In view of the structure of $\Phi(k)$, we have that

$$\Pi_{k_2, k_1}^\Phi = \begin{bmatrix} \Pi_{k_2, k_1}^A & \Omega_{k_2, k_1} \\ 0 & \Pi_{k_2, k_1}^B \end{bmatrix},$$

where $\Omega_{k_2, k_1} = \sum_{k=k_1}^{k_2} \Pi_{k_2, k+1}^A D \Pi_{k-1, k_1}^B$. Since Π_{k_2, k_1}^A is bounded as $k_2 \rightarrow \infty$, we can define

$$G_1 = \sup_{k_2 \geq k_1 \geq k_I} \|\Pi_{k_2, k_1}^A D\|_2.$$

Because $B_i(k)$ will converge to 0 as $k \rightarrow \infty$ according to (8.24), we can easily arrive at that $\|\Pi_{k_2, k_1}^B\|_2$ converges to 0 very fast as $k_2 \rightarrow \infty$ (even faster than an

exponential rate). Moreover, $\sum_{k_2=k_1}^{\infty} \|\Pi_{k_2,k_1}^B\|_2 < +\infty$ for any given k_1 . Therefore, for any arbitrarily small number $\varepsilon > 0$, there exists $N(\varepsilon) \in \mathbb{R}^+$ such that

$$\sum_{k_2=N(\varepsilon)}^{\infty} \|\Pi_{k_2,k_1}^B\|_2 < \frac{\varepsilon}{\sqrt{2}G_1}.$$

Hereafter, we define $e_{m,i,j} \triangleq e_{m,i} - e_{m,j}$. Then, it follows that for any given $k_1 \geq k_I$,

$$\lim_{k_2 \rightarrow \infty} \|\Omega_{k_2,k_1}\|_2 = \lim_{k_2 \rightarrow \infty} \left\| \sum_{k=k_1}^{k_2} \Pi_{k_2,k+1}^A D \Pi_{k-1,k_1}^B \right\|_2 \leq G_1 \sum_{k=k_1}^{\infty} \|\Pi_{k-1,k_1}^B\|_2 < +\infty, \quad (8.33)$$

and $\forall i, j \in \mathcal{V}$,

$$\begin{aligned} \lim_{k_2 \rightarrow \infty} \|e'_{n,i,j} \Omega_{k_2,k_1}\|_2 &= \lim_{k_2 \rightarrow \infty} \left\| e'_{n,i,j} \sum_{k=k_1}^{k_2} \Pi_{k_2,k+1}^A D \Pi_{k-1,k_1}^B \right\|_2 \\ &\leq \lim_{k_2 \rightarrow \infty} \left\| e'_{n,i,j} \sum_{k=k_1}^{N(\varepsilon)-1} \Pi_{k_2,k+1}^A D \Pi_{k-1,k_1}^B \right\|_2 \\ &\quad + \lim_{k_2 \rightarrow \infty} \left\| e'_{n,i,j} \sum_{k=N(\varepsilon)}^{k_2} \Pi_{k_2,k+1}^A D \Pi_{k-1,k_1}^B \right\|_2 \\ &\leq \|e'_{n,i,j}\|_2 G_1 \sum_{k=N(\varepsilon)}^{\infty} \|\Pi_{k-1,k_1}^B\|_2 \\ &< \varepsilon. \end{aligned} \quad (8.34)$$

In the above, the second inequality is due to (8.30). (8.33) and (8.34) imply that Ω_{k_2,k_1} converges to a bounded identical row matrix as $k_2 \rightarrow \infty$ with the converged row denoted by $r_2(k_1)$. Then we have

$$\lim_{k_2 \rightarrow \infty} \Pi_{k_2,k_1}^\Phi = [\mathbf{1}'_n \quad \mathbf{0}'_{n(\bar{d}+1)}]' [r_1(k_1) \quad r_2(k_1)], \quad (8.35)$$

where $r_1(k_1)$ is as defined in (8.30). According to (8.31), we write $z(k)$ in the following way

$$z(k) = \Pi_{k-1,k_I}^\Phi z(k_I) + \sum_{l=k_I+1}^k c(l) \Pi_{k-1,l}^\Phi \Gamma W(l). \quad (8.36)$$

Because the matrix $\Pi_{k-1,j}^\Phi \Gamma$ is uniformly bounded, $W(j)$ satisfies Assumption 8.1 and $c(k)$ satisfies $\sum_{k=0}^{\infty} c^2(k) < +\infty$, by recalling Khintchine-Kolmogorov convergence theorem [19], we know that the second term on the right-hand side of (8.36) will converge to a random vector almost surely as $k \rightarrow \infty$. From (8.35), we see that the first term on the right-hand side of (8.36) converges to $[\mathbf{1}'_n \quad \mathbf{0}'_{n(\bar{d}+1)}]' [r_1(k_I) \quad r_2(k_I)] z(k_I)$ as $k \rightarrow \infty$. Therefore, we have proved that $z(k)$ converge to a random vector almost surely. Next, we can prove that the first n entries of $z(k)$ converge to a same random variable with probability one. According to (8.27), for any arbitrarily small number $\varepsilon > 0$, there must exist $N(\varepsilon)$ such that

$$\sum_{k=N(\varepsilon)}^{\infty} c^2(k) < \frac{\varepsilon}{2G_2}, \quad (8.37)$$

where $G_2 \triangleq \sup_{k_2 \geq k_1 \geq k_I} (\|\Pi_{k_2,k_1}^\Phi \Gamma\|_2^2 \|Var(W(k_1))\|_2) < +\infty$. For all $i, j \in \mathcal{V}$, since $z(k_I)$ is independent of $W(k)$, $\forall k > k_I$, we have

$$\begin{aligned} \lim_{k \rightarrow \infty} E \left| e'_{n(\bar{d}+1),i,j} z(k) \right|^2 &= \lim_{k \rightarrow \infty} E \left| e'_{n(\bar{d}+1),i,j} \Pi_{k-1,k_I}^\Phi z(k_I) \right|^2 \\ &\quad + \lim_{k \rightarrow \infty} Var \left(e'_{n(\bar{d}+1),i,j} \sum_{l=k_I+1}^{N(\varepsilon)-1} c(l) \Pi_{k-1,l}^\Phi \Gamma W(l) \right) \\ &\quad + \lim_{k \rightarrow \infty} Var \left(e'_{n(\bar{d}+1),i,j} \sum_{l=N(\varepsilon)}^k c(l) \Pi_{k-1,l}^\Phi \Gamma W(l) \right) \\ &\leq \lim_{k \rightarrow \infty} e'_{n(\bar{d}+1),i,j} \sum_{l=N(\varepsilon)}^k c^2(l) \Pi_{k-1,l}^\Phi \Gamma Var(W(l)) (\Pi_{k-1,l}^\Phi \Gamma)' \\ &\quad \cdot e_{n(\bar{d}+1),i,j} \\ &\leq 2G_2 \sum_{l=N(\varepsilon)}^{\infty} c^2(l) \\ &< \varepsilon. \end{aligned}$$

In the above, the second inequality is due to the definition of G_2 and the last inequality follows from (8.37). The arbitrariness of ε implies that $\forall i, j \in \mathcal{V}$,

$$\lim_{k \rightarrow \infty} E \left| e'_{n(\bar{d}+1),i,j} z(k) \right|^2 = 0. \quad (8.38)$$

Since we know that $e'_{n(\bar{d}+1),i,j}z(k)$ converges to a random variable almost surely, by recalling [126], we can say that $\forall i, j \in \mathcal{V}$, $e'_{n(\bar{d}+1),i,j}z(k)$ converges to 0 almost surely. Then, by the definition of $z(k)$, strong consensus is proved.

‘Necessity’. The proof is divided into the proofs of the following:

1) The quotient graph contains a spanning tree.

2) $\sum_{k=0}^{\infty} c(k) = +\infty$.

3) $\sum_{k=0}^{\infty} c^2(k) < +\infty$.

1). By recalling [101], the MAS with $W(k) = 0$, $k \in \bar{\mathbb{Z}}^+$ cannot reach consensus If the quotient graph does not contain a spanning tree.

2). By contradiction, we assume that $\sum_{k=0}^{\infty} c(k) < +\infty$, which can imply that $\sum_{k=0}^{\infty} c^2(k) < +\infty$. We can still write $A(k) = TA_1(k)T^{-1}$, where $A_1(k)$ is as defined in (8.29). According to Lemma 8.3, $\Pi_{k_2,k_1}^{A_1}$ cannot converge to a matrix with the last $n-1$ rows equal to 0. So there exists $i \in \mathcal{V}$ such that

$$e'_{n,j} \lim_{k_2 \rightarrow \infty} \Pi_{k_2,k_1}^{A_1} e_{n,i} \neq 0, \quad (8.39)$$

for some $j \in \mathcal{V} \setminus \{1\}$. Then we consider system (8.9). By taking initial condition $X(k_I) = Te_{n,i}$ with i satisfying (8.39), $W_c(k_I - l) = 0$, $l = 0, \dots, \bar{d} - 1$ and noise input $W(k) = 0$, $k \geq k_I$, we have

$$\lim_{k \rightarrow \infty} X(k) = \lim_{k \rightarrow \infty} \Pi_{k,k_I}^A X(k_I) = T \lim_{k \rightarrow \infty} \Pi_{k,k_I}^{A_1} e_{n,i}.$$

According to (8.39) and the definition of T , all the elements of $X(k)$ cannot converge to a common value. Hence, $\sum_{k=0}^{\infty} c(k) = +\infty$ is necessary.

3). By contradiction, we assume that $\sum_{k=0}^{\infty} c^2(k) = +\infty$. We write the system (8.8) in the following augmented form

$$\bar{X}(k+1) = \Psi(k)\bar{X}(k) + c(k)\Theta W(k), \quad (8.40)$$

where $\bar{X}(k) = \text{col}\{X(k), X(k-1), \dots, X(k-\bar{d})\}$ and

$$\Psi(k) = \begin{bmatrix} I - c(k)L_0(k) & -c(k)L_1(k) & \cdots & -c(k)L_{\bar{d}}(k) \\ & I & & 0 \\ & & \ddots & \vdots \\ & & & I & 0 \end{bmatrix}, \Theta = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then, $\bar{X}(k)$ is given by

$$\bar{X}(k) = \Pi_{k-1,0}^\Psi \bar{X}(0) + \sum_{l=0}^{k-1} \Pi_{k-1,l+1}^\Psi c(l)\Theta W(l).$$

Since there exists a random variable x^* such that

$$\lim_{k \rightarrow \infty} \bar{X}(k) = \mathbf{1}_{n(\bar{d}+1)} x^*,$$

for all $\bar{X}(0)$ and $W(k)$ satisfying (8.6), by choosing $\{W(k) = 0, \forall k \in \bar{\mathbb{Z}}^+\}$, it is clear that Π_{k_2,k_1}^Ψ must converge to an identical row matrix as $k_2 \rightarrow \infty$, so we write

$$\lim_{k_2 \rightarrow \infty} \Pi_{k_2,k_1}^\Psi = \mathbf{1}_{n(\bar{d}+1)} [q_0(k_1) \ q_1(k_1) \ \cdots \ q_{\bar{d}}(k_1)], \quad (8.41)$$

where $q_l(k_1)$, $l = 0, 1, \dots, \bar{d}$ are n dimensional row vectors. On the other hand, since $\Psi \mathbf{1}_{n(\bar{d}+1)} = \mathbf{1}_{n(\bar{d}+1)}$ and therefore $\Pi_{k_2,k_1}^\Psi \mathbf{1}_{n(\bar{d}+1)} = \mathbf{1}_{n(\bar{d}+1)}$, we have

$$\sum_{l=0}^{\bar{d}} q_l(k_1) \mathbf{1}_n = \mathbf{1}_n, \quad \forall k_1 \in \bar{\mathbb{Z}}^+. \quad (8.42)$$

It means that there exists $0 \leq l \leq \bar{d}$ such that $q_l(k_1) \neq 0$. Next, we shall prove that

$$\lim_{k \rightarrow \infty} q_0(k) \neq 0.$$

By contradiction, we assume that for any small number ε , there exists $N(\varepsilon)$ such

that

$$\|q_0(k)\|_\infty < \varepsilon, \quad \forall k > N(\varepsilon).$$

Before proceeding further, we need to make clear that $c(k)$ is uniformly bounded. Actually, if $c(k)$ diverges, Π_{k_2, k_1}^Ψ does not exist as $k_2 \rightarrow \infty$. So we denote $\bar{c} = \sup_{k \geq 0} c(k)$. According to the definition of $q_i(k)$ in (8.41), by simple calculation, we have

$$\begin{aligned} q_0(k) &= q_0(k+1) - \sum_{l=0}^{\bar{d}} c(k+l) q_0(k+l+1) L_l(k+l), \\ q_i(k) &= - \sum_{l=i}^{\bar{d}} c(k+l-i) q_0(k+l+1-i) L_l(k+l-i), \quad i = 1, \dots, \bar{d}. \end{aligned}$$

Then, for all $k \geq N(\varepsilon)$,

$$\begin{aligned} \|q_0(k)\|_\infty &= \left\| q_0(k+1) - \sum_{l=0}^{\bar{d}} c(k+l) q_0(k+l+1) L_l(k+l) \right\|_\infty \\ &\leq \varepsilon (1 + \bar{c}(\bar{d}+1) \|L\|_\infty), \end{aligned} \quad (8.43)$$

$$\begin{aligned} \|q_i(k)\|_\infty &= \left\| \sum_{l=i}^{\bar{d}} c(k+l-i) q_0(k+l+1-i) L_l(k+l-i) \right\|_\infty \\ &\leq \varepsilon \bar{c}(\bar{d}-i+1) \|L\|_\infty. \end{aligned} \quad (8.44)$$

On the other hand, according to (8.42), $\sum_{l=0}^{\bar{d}} \|q_l(k)\|_\infty \geq 1$, which contradicts (8.43)-(8.44) by choosing

$$\varepsilon \leq \frac{2\bar{c}\|L\|_\infty}{(\bar{d}+1)(\bar{d}+2)}.$$

Hence, there exist ε and $N_1(\varepsilon)$ such that for all $k \geq N_1(\varepsilon)$, $\|q_0(k)\|_\infty > \varepsilon$. Denote

$$\xi = \inf_{k \geq N_1(\varepsilon)} \|q_0(k)\Xi\|_2, \quad (8.45)$$

where Ξ is defined in (8.7). Note that $\forall i \in \mathcal{V}$, $\deg_{in}(i) > 0$, the row rank of Ξ is full, which implies that $\xi > 0$. Since $\sum_{k=0}^{\infty} c^2(k) = +\infty$, for any large number K ,

there exists N_2 such that

$$\sum_{k=N_1(\varepsilon)}^{N_2} c^2(k) > \frac{K}{n(\bar{d}+1)\xi^2}. \quad (8.46)$$

Then, by setting $\bar{X}(0) = 0$ and $\{W(k), k \in \mathbb{Z}^+\}$ the white noise process with $\text{Var}(W(k)) = \Xi\Xi' > 0$, we have

$$\begin{aligned} E\bar{X}'(k)\bar{X}(k) &= E \left[\sum_{l=1}^{k-1} (\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l))' (\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l)) \right] \\ &= \sum_{l=1}^{k-1} E \left[(\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l))' (\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l)) \right] \\ &\geq \sum_{l=N_1(\varepsilon)}^{N_2} E \left[(\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l))' (\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l)) \right], \quad \forall k > N_2, \end{aligned}$$

and,

$$\begin{aligned} \lim_{k \rightarrow \infty} E\bar{X}'(k)\bar{X}(k) &\geq \lim_{k \rightarrow \infty} \sum_{l=N_1(\varepsilon)}^{N_2} E \left[(\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l))' (\Pi_{k-1,l+1}^\Psi c(l)\Theta W(l)) \right] \\ &= \sum_{l=N_1(\varepsilon)}^{N_2} E \left[(\mathbf{1}_{n(\bar{d}+1)} q_0(l+1)c(l)W(l))' (\mathbf{1}_{n(\bar{d}+1)} q_0(l+1)c(l)W(l)) \right] \\ &= n(\bar{d}+1) \sum_{l=N_1(\varepsilon)}^{N_2} c^2(l)q_0(l+1)\Xi\Xi'q_0'(l+1) \\ &\geq n(\bar{d}+1)\xi^2 \sum_{l=N_1(\varepsilon)}^{N_2} c^2(l) \\ &> K. \end{aligned}$$

The second to the last inequality comes from (8.45) and the last inequality from (8.46). The arbitrariness of K implies that

$$\lim_{k \rightarrow \infty} E\bar{X}'(k)\bar{X}(k) = +\infty. \quad (8.47)$$

On the other hand, if the system reaches strong consensus, $x_i(k)$ must converge to

a random variable almost surely as $k \rightarrow \infty$. By recalling Kolmogorov three series theorem [19], we see that $x_i(k)$ converges to a random variable almost surely only if

$$\lim_{k \rightarrow \infty} \sum_{i \in \mathcal{V}} E x_i^2(k) < +\infty. \quad (8.48)$$

However, according to (8.47), it is noted that

$$\lim_{k \rightarrow \infty} \sum_{i \in \mathcal{V}} E x_i^2(k) = \frac{1}{\bar{d} + 1} \lim_{k \rightarrow \infty} E \bar{X}'(k) \bar{X}(k) = +\infty,$$

which contradicts (8.48). Hence, $\sum_{k=1}^{\infty} c^2(k) = +\infty$ does not hold. $\square$

Theorem 8.2. (Leader-follower case) *Consider the MAS (8.1)-(8.2) with leader-follower topology and interaction protocol (8.5). For any noises satisfying Assumption 8.1 and topology with at least one vertex i satisfying $\deg_{in}(i) = 0$, the strong consensus is achieved if and only if the quotient graph contains a spanning tree and the consensus gain function satisfies (8.25).*

Proof. ‘**Sufficiency**’. Since the graph contains a spanning tree, there exists only one vertex whose in-degree is 0. Without loss of generality, we assume that $\deg_{in}(n) = 0$, which implies that

$$x_n(k+1) = x_n(k), \quad \forall k \in \bar{\mathbb{Z}}^+.$$

Define $\tilde{x}_i(k) = x_i(k) - x_n(k)$, $i = 1, \dots, n-1$. By applying the protocol (8.5), we have the following system

$$\tilde{X}(k+1) = \tilde{X}(k) - c(k) \sum_{l=0}^{\bar{d}} L_l(k) \tilde{X}(k-l) + c(k) W(k), \quad (8.49)$$

where $\tilde{X}(k) = \text{col}\{\tilde{x}_1(k), \dots, \tilde{x}_{n-1}(k)\}$; $W(k) = \text{col}\{w_1(k), \dots, w_{n-1}(k)\}$ with $w_i(k)$

as defined in (8.6); $L_l(k)$, $l = 0, \dots, \bar{d}$ are $(n-1) \times (n-1)$ matrices with

$$L_l^{i,j}(k) \triangleq \begin{cases} -a_{i,j}\delta_{d_{i,j}(k),l}, & i \neq j, \\ \sum_{m \in \mathcal{V} \setminus \{i\}} a_{i,m}\delta_{t_{i,m}(k),l}, & i = j. \end{cases}$$

The strong consensus is equivalent to that

$$\lim_{k \rightarrow \infty} \tilde{X}(k) = 0, \text{ a.s.}$$

System (8.49) can be written in the form of (8.9) with $X_1(k)$ replaced by $\tilde{X}(k)$ and initial condition $A(k_I - l)$, $l = 1, \dots, \bar{d}$ satisfying (8.18), which can be further written in the augmented form (8.31) with $X_1(k)$ replaced by $\tilde{X}(k)$ in $z(k)$. Denote $L_{(n-1)}$ the square matrix by deleting the last row and last column of L . It is noted that $\sum_{l=0}^{\bar{d}} L_l(k) = L_{(n-1)}$. Then, according to Lemma 8.1

$$\|B_i(k)\|_2 = \left\| \sum_{l=i}^{\bar{d}} c(k)L_l(k) + \Delta_{B,i}(k) \right\|_2 \leq c(k)\|L_{(n-1)}\|_2 + C\rho_{\bar{d}}(k),$$

where $C > 0$ is as defined in (8.17), $\rho_{\bar{d}}(k)$ is as defined in (8.14). Since $\lim_{k \rightarrow \infty} c(k) = 0$, for any small number ε , there exists $N(\varepsilon)$ such that for all $i = 1, \dots, \bar{d}$

$$\|B_i(k)\|_\infty < \varepsilon, \quad \forall k > N(\varepsilon). \quad (8.50)$$

Define

$$H = \begin{cases} \text{diag}\{I, \varepsilon^{\frac{1}{\bar{d}}}I, \varepsilon^{\frac{2}{\bar{d}}}I, \dots, \varepsilon^{\frac{\bar{d}-1}{\bar{d}}}I\}, & \bar{d} > 1, \\ I, & \bar{d} = 1, \end{cases}$$

and

$$\bar{H} = \begin{bmatrix} I & 0 \\ 0 & H \end{bmatrix}, \quad \tilde{\Phi}(k) = \bar{H}\Phi(k)\bar{H}^{-1} = \begin{bmatrix} A(k) & D \\ 0 & \tilde{B}(k) \end{bmatrix},$$

where $\tilde{B}(k) = HB(k)H^{-1}$, $\Phi(k)$ is as defined in (8.32). It is easy to check that $\forall k > N(\varepsilon)$,

$$\|\tilde{B}(k)\|_\infty \leq \frac{\varepsilon^{\frac{1}{\bar{d}}}}{1 - \varepsilon^{\frac{1}{\bar{d}}}}. \quad (8.51)$$

On the other hand, $\Pi_{k_2, k_1}^{\tilde{\Phi}}$ can be written as follows

$$\Pi_{k_2, k_1}^{\tilde{\Phi}} = \begin{bmatrix} \Pi_{k_2, k_1}^A & \tilde{\Omega}_{k_2, k_1} \\ 0 & \Pi_{k_2, k_1}^{\tilde{B}} \end{bmatrix},$$

where $\tilde{\Omega}_{k_2, k_1} = \sum_{k=k_1}^{k_2} \Pi_{k_2, k+1}^A D \Pi_{k-1, k_1}^{\tilde{B}}$. According to Lemma A.1, the real component of any eigenvalue of $L_{(n-1)}$ is positive. Then, by Lemma 8.1, there exist $C_2, C_3, C_4 > 0$ such that

$$\|\Pi_{k_2, k_1}^A\|_2 \leq C_2 \prod_{k=k_1}^{k_2} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k)), \quad k_2 \geq k_1 \geq k_I. \quad (8.52)$$

It is noted that $\|\tilde{B}(k)\|_2 \leq \sqrt{dn} \|\tilde{B}(k)\|_\infty$. Define

$$k'_I = k_I \vee N \left(\left(3\sqrt{dn} + 3 \right)^{-\bar{d}} \right) \vee s \left(\frac{1}{2C_3} \right),$$

where $N(\cdot)$ is as defined in (8.50), $s(\cdot)$ is as defined in (8.16). Then, it follows $\forall k \geq k'_I$,

$$1 - C_3 c(k) + C_4 \rho(k) \geq 1 - C_3 c(k) \geq \frac{1}{2}, \quad \|\tilde{B}(k)\|_2 \leq \frac{1}{3}. \quad (8.53)$$

By considering (8.51) and (8.52), we have $\forall k_1 \geq k'_I$,

$$\begin{aligned} \left\| \tilde{\Omega}_{k_2, k_1} \right\|_2 &\leq \sum_{k=k_1}^{k_2} \left\| \Pi_{k_2, k+1}^A \right\|_2 \|D\|_2 \left\| \Pi_{k-1, k_1}^{\tilde{B}} \right\|_2 \\ &\leq \sum_{k=k_1}^{k_2} C_2 \left(\frac{1}{3} \right)^{k-k_1} \prod_{l=k+1}^{k_2} (1 - C_3 c(l) + C_4 \rho_{\bar{d}}(l)) \\ &\leq C_2 \prod_{k=k_1+1}^{k_2} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k)) \sum_{k=k_1}^{k_2} \left[\left(\frac{1}{3} \right)^{k-k_1} \right. \\ &\quad \cdot \left. \prod_{l=k_1+1}^k (1 - C_3 c(l) + C_4 \rho_{\bar{d}}(l))^{-1} \right] \\ &\leq C_2 \prod_{k=k_1+1}^{k_2} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k)) \sum_{k=k_1}^{k_2} \left(\frac{2}{3} \right)^{k-k_1} \end{aligned}$$

$$\leq 3C_2 \prod_{k=k_1+1}^{k_2} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k)). \quad (8.54)$$

The second and fourth inequality come from (8.53). According to (8.31), we write $\tilde{X}(k+1)$ in the following form

$$\begin{aligned} \tilde{X}(k+1) &= [I \ 0 \ \cdots \ 0] \Pi_{k,k'_I}^\Phi z(k'_I) + [I \ 0 \ \cdots \ 0] \sum_{l=k'_I+1}^{k+1} c(l) \Pi_{k,l}^\Phi \Gamma W(l) \\ &= [\Pi_{k,k'_I}^A \ \tilde{\Omega}_{k,k'_I}] \bar{H} z(k'_I) + \sum_{l=k'_I+1}^{k+1} c(l) \tilde{\Omega}_{k,l} D^T W(l), \end{aligned} \quad (8.55)$$

and,

$$\begin{aligned} \left\| E\tilde{X}(k+1)\tilde{X}'(k+1) \right\|_2 &\leq \left\| [\Pi_{k,k'_I}^A \ \tilde{\Omega}_{k,k'_I}] \bar{H} \text{Var}(z(k'_I)) \bar{H}' \begin{bmatrix} (\Pi_{k,k'_I}^A)' \\ \tilde{\Omega}_{k,k'_I}' \end{bmatrix} \right\|_2 \\ &\quad + \left\| \sum_{l=k'_I+1}^{k+1} c^2(l) \tilde{\Omega}_{k,l} D' \text{Var}(W(l)) D \tilde{\Omega}_{k,l}^T \right\|_2. \end{aligned} \quad (8.56)$$

This is due to that $z(k'_I)$ is independent of the noises $W(k)$, $k > k'_I$. Since $c(k)$ satisfies (8.25), in view of the proof of part 1) of Lemma 8.3, we can see that $\forall k_1 \geq 0$,

$$\lim_{k_2 \rightarrow \infty} \prod_{k=k_1}^{k_2} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k)) = 0.$$

Note that k'_I is finite and therefore $\text{Var}(z(k'_I))$ is bounded, so, according to (8.52) and (8.54), the first term on the right-hand side of (8.56) goes to 0 as $k \rightarrow \infty$. By denoting $\sup_{k \geq 0} \|\text{Var}(W(k))\|_2 = \bar{W}$, we have

$$\begin{aligned} &\lim_{k \rightarrow \infty} \left\| \sum_{l=k'_I+1}^{k+1} c^2(l) \tilde{\Omega}_{k,l} D' \text{Var}(W(l)) D \tilde{\Omega}_{k,l}' \right\|_2 \\ &\leq \lim_{k \rightarrow \infty} \sum_{l=k'_I+1}^{k+1} c^2(l) \left\| \tilde{\Omega}_{k,l} \right\|_2^2 \|D\|_2^2 \|\text{Var}(W(l))\|_2 \end{aligned}$$

$$\begin{aligned}
&\leq \lim_{k \rightarrow \infty} 9C_2^2 \bar{W} \sum_{l=k_I'+1}^{k+1} c^2(l) \prod_{k=l+1}^{k+1} (1 - C_3 c(k) + C_4 \rho_{\bar{d}}(k))^2 \\
&= \lim_{k \rightarrow \infty} \frac{9C_2^2 \bar{W} c^2(k)}{2(C_3 c(k) - C_4 \rho_{\bar{d}}(k)) - (C_3 c(k) - C_4 \rho_{\bar{d}}(k))^2} \\
&= 0.
\end{aligned} \tag{8.57}$$

The second last equality comes from Lemma 8.4. (8.57) implies that the second term on the right-hand side of (8.56) goes to 0 as $k \rightarrow \infty$. Hence, we have proved that $\lim_{k \rightarrow \infty} E\tilde{X}(k)\tilde{X}'(k) = 0$. By recalling Khintchine-Kolmogorov convergence theorem [19], we know that $\tilde{X}(k)$ converges to 0 almost surely, which completes the proof of the sufficiency.

‘Necessity’. The necessity of ‘the graph contains a spanning tree’ and $\sum_{k=1}^{\infty} c(k) = +\infty$ can be proved by the same argument as in the proof of the necessity of Theorem 8.1.

Next, we prove $\lim_{k \rightarrow \infty} c(k) = 0$ by contradiction. Assume that there exists $\varepsilon > 0$ such that for any large number K , there exists $k > K$ satisfying $c(k) \geq \varepsilon$. For equation (8.55), by setting $z(k_I) = 0$ and $Var(W(k)) = \bar{W}$, we have

$$Var\left(\tilde{X}(k+1)\right) \geq Var\left(\sum_{l=k_I+1}^{k-1} c(l)\tilde{\Omega}_{k-1,l}D'W(l)\right) + c^2(k)\bar{W} \geq \varepsilon^2\bar{W} \neq 0.$$

The arbitrariness of K implies that $\lim_{k \rightarrow \infty} Var(\tilde{X}(k)) \neq 0$, which contradicts the strong consensus. So, $\lim_{k \rightarrow \infty} c(k) = 0$. $\square$

So far, we have provided necessary and sufficient conditions for the strong consensus of system (8.1)-(8.2) with protocol (8.5). Next, we shall show that the conditions are also necessary and sufficient for mean square consensus.

Theorem 8.3. (Non-leader-follower case) *Consider system (8.1)-(8.2) with interaction protocol (8.5). For any noises satisfying Assumption 8.1 and any topology satisfying $\deg_{in}(i) > 0, \forall i \in \mathcal{V}$, the mean square consensus is achieved if and only if the quotient graph contains a spanning tree and the consensus gain function satisfies (8.27).*

Proof. ‘Sufficiency’. According to (8.38), we have obtained that the second order moment of the difference between $x_i(k)$ and $x_j(k)$ will tend to 0 as $k \rightarrow \infty$ for all $i, j \in \mathcal{V}$, i.e.

$$\lim_{k \rightarrow \infty} E |x_i(k) - x_j(k)|^2 = \lim_{k \rightarrow \infty} E \left| e'_{n(\bar{d}+1), i, j} z(k) \right|^2 = 0, \quad \forall i, j \in \mathcal{V}. \quad (8.58)$$

On the other hand, there exists a random vector $z^* \triangleq \text{col}\{z_1^*, z_2^*, \dots, z_{n(\bar{d}+1)}^*\}$ such that

$$\lim_{k \rightarrow \infty} E \|z(k) - z^*\|_2^2 = 0. \quad (8.59)$$

By combining (8.58) and (8.59), $\forall i \in \mathcal{V}$, we have

$$\begin{aligned} \lim_{k \rightarrow \infty} E |x_i(k) - z_1^*|^2 &= \lim_{k \rightarrow \infty} E |(x_i(k) - x_1(k)) + (x_1(k) - z_1^*)|^2 \\ &\leq 2 \lim_{k \rightarrow \infty} E |x_i(k) - x_1(k)|^2 + 2 \lim_{k \rightarrow \infty} E |x_1(k) - z_1^*|^2 \\ &= 0, \end{aligned}$$

which implies that $\lim_{k \rightarrow \infty} E |x_i(k) - z_1^*|^2 = 0$, $i = 1, \dots, n$. The mean square consensus is obtained.

‘Necessity’. If the system reaches mean square consensus, the quotient graph must contain a spanning tree and $\sum_{k=1}^{\infty} c(k) = +\infty$ holds. This can be proved by the same argument as in the proof of the necessity of Theorem 8.1. For the condition $\sum_{k=1}^{\infty} c^2(k) < +\infty$, we can prove it by contradiction. We still consider system (8.40) with initial value $\bar{X}(0) = 0$ and white noise process $\{W(k), k \in \mathbb{Z}^+\}$ with $\text{Var}(W(k)) = \Xi\Xi' > 0$. According to (8.47), if $\sum_{k=0}^{\infty} c^2(k) = +\infty$, then $\lim_{k \rightarrow \infty} \sum_{i=1}^n E x_i^2(k) = +\infty$. On the other hand, if the mean square consensus is reached, $x_i(k)$ must converge in the mean square sense to a random variable x^* with a finite second order moment. By dominated convergence theorem [126], $\lim_{k \rightarrow \infty} E x_i^2(k) = E(x^*)^2 < +\infty$, and therefore $\lim_{k \rightarrow \infty} \sum_{i=1}^n E x_i^2(k) = nE(x^*)^2 < +\infty$, which leads to a contradiction. The proof is completed. $\square$

Theorem 8.4. (Leader-follower case) *Consider system (8.1)-(8.2) with interaction protocol (8.5). For any noises satisfying Assumption 8.1 and topology with at least one vertex i satisfying $\deg_{in}(i) = 0$, the mean square consensus is achieved if and only if the quotient graph contains a spanning tree and the consensus gain function satisfies (8.25).*

Proof. The proof is similar to that of proof of Theorem 8.3 by taking into account of the proof of Theorem 8.2. $\square$

Remark 8.2. *Theorems 8.1-8.4 present necessary and sufficient conditions for strong consensus and mean square consensus, respectively. Note that the techniques proposed in [48], [66] and [53] for the delay-free case cannot be directly applied to our problems. For the delay-free case, the whole system can be written in the following compact form*

$$X(k+1) = (I - c(k)L)X(k) + c(k)W(k) \quad (8.60)$$

whose stability is much easier to analyze than that of (8.8). Note that the standard state augmentation approach cannot directly be applied as it will result in an augmented system whose state matrix is not given in terms of the Laplacian matrix L of the quotient graph, which will make the stability analysis difficult. By introducing the auxiliary system (8.9), a new approach has been developed for consensus analysis by establishing the convergence properties of $\Pi_{i,j}^\Phi$. It is noted that when delays disappear, the matrices $B_i(k)$, $i = 1, 2, \dots, \bar{d}$ become 0 and the results reduce to the existing works [48], [66] and [53].

Remark 8.3. *For the leader-follower case, due to that all the followers are required to converge to the state of the leader, the condition on the gain function can be weaker than that for the non-leader-follower case. Note that the condition (8.25) in Theorem 8.2 and 8.4 is shown to be a sufficient condition for ensuring mean square weak consensus in [65]. In this chapter, we have established that the condition (8.25) together with the connectivity condition are necessary and sufficient for MASs with a leader-follower structure to reach strong and mean square consensus. It is worth mentioning that our result can be extended to improving the result in [75].*

8.2.2 Dynamically Switching Topologies

We consider the case where the topology is dynamically switching. The graph $\mathcal{G}(k)$ is thus time dependent with the vertex set $\mathcal{V}$ and the edge set $\mathcal{E}(k)$.

In most practical MASs, topology cannot be always fixed. For example, in flocking or formation problem, communications may depend on the distance between each pair of agents which is time-varying. Another example is that for a large scale wireless network, time synchronization is impossible [38], which means that the actuators cannot act simultaneously. In particular, the gossip algorithm [5], [26] can be seen as an asynchronous protocol. In that case, the communication graph is a dynamically switching graph.

Since the Laplacian matrix of the quotient graph becomes a time-varying matrix, the analysis of the convergence is much more difficult. Some general conditions such as that the communication graph is connected for an infinite number of times should be introduced to guarantee that the agents are connected in the long term. For noise free systems, there are some existing works on the consensus problem under switching topologies [51], [101], [12]. By leveraging on these existing works, the results in the last section can be extended to the case of dynamically changing topologies.

The distributed interaction protocol for $\{\mathcal{G}(k), k \in \bar{\mathbb{Z}}^+\}$ is given as follows:

$$u_i(k) = c(k) \sum_{j \in \mathcal{V}} a_{i,j}(k) (y_{i,j}(k) - x_i(k - t_{i,j}(k))), \quad (8.61)$$

where $a_{i,j}(k) \in \mathcal{C} \cup \{0\}$.

Before investigating the consensus problem, we introduce the notion of union graph. Given a sequence of graphs, the union graph is a graph whose vertex and edge sets are respectively the union of the vertex and edge sets of these graphs. For example, we denote $\mathcal{G}(k, h)$ the union graph of $\{\mathcal{G}(i), k \leq i < k + h\}$, then we have

$$\mathcal{V}(\mathcal{G}(k, h)) = \bigcup_{i=k}^{k+h-1} \mathcal{V}(\mathcal{G}(i)), \quad \mathcal{E}(\mathcal{G}(k, h)) = \bigcup_{i=k}^{k+h-1} \mathcal{E}(\mathcal{G}(i)).$$

To guarantee that the graph is connected in the long term, we need the following assumption:

Assumption 8.2. *The quotient graph $\mathcal{G}_Q(k)$ is balanced at any time instant and there exists a constant $h \in \mathbb{Z}^+$ such that for any $k \in \bar{\mathbb{Z}}^+$, the union of the quotient graphs, $\mathcal{G}_Q(k, h)$, contains a spanning tree.*

Assumption 8.2 which guarantees that the agents are connected for an infinite number of times, is very standard [3], [12]. Then we have the following result.

Theorem 8.5. *Consider the MAS (8.1)-(8.2) with dynamically switching topologies. Given distributed protocol (8.61), under Assumptions 8.1 and 8.2, both strong consensus and mean square consensus are achieved if $c(k)$ satisfies (8.27) and*

$$0 < \inf_{k \geq 0} \frac{c(k)}{c(k+1)} \leq \sup_{k \geq 0} \frac{c(k)}{c(k+1)} < +\infty. \quad (8.62)$$

Proof. We write the closed-loop MAS of (8.1)-(8.2) with interaction protocol (8.61) in the following compact form

$$X(k+1) = X(k) - c(k) \sum_{l=0}^{\bar{d}} L_l(k) X(k-l) + c(k) W(k), \quad (8.63)$$

where $X(k)$ and $W(k)$ are as defined in (8.6),

$$L_l^{i,j}(k) \triangleq \begin{cases} -a_{i,j}(k) \delta_{d_{i,j}(k),l}, & i \neq j, \\ \sum_{m \in \mathcal{V} \setminus \{i\}} a_{i,m}(k) \delta_{t_{i,m}(k),l}, & i = j. \end{cases} \quad (8.64)$$

Note that $\sum_{l=0}^{\bar{d}} L_l(k) \triangleq L(k)$ is the Laplacian matrix of the quotient graph $\mathcal{G}_Q(k)$. Since $a_{i,j}(k)$ belongs to a compact set, $L(k)$ must be bounded. We can write system (8.63) into (8.9) with starting time instant $k_I = k_0 + \bar{d}$, where k_0 is as defined in (8.17), and initial conditions on $A(k_I - l)$, $l = 1, \dots, \bar{d}$ satisfying (8.18). According to Lemma 8.1 we know that $\forall k \geq k_I$

$$\|A(k) - I + c(k)L(k)\|_2 \leq C\rho_{\bar{d}}(k), \quad \left\| B_i(k) - c(k) \sum_{j=i}^{\bar{d}} L_j(k) \right\|_2 \leq C\rho_{\bar{d}}(k),$$

where C is as defined in (8.17) with $\|L\|_\infty$ replaced by $\sup_{k \in \mathbb{Z}^+} \|L(k)\|_\infty$. By applying similar arguments as in proving (8.23), we have that there exists C_1 such that $\forall k \geq k_I$,

$$\left\| \Pi_{k+h, k+1}^A - I + \sum_{l=k+1}^{k+h} c(l)L(l) \right\|_2 \leq C_1 \rho_{\bar{d}+h}(k), \quad (8.65)$$

Due to (8.62) and Assumption 8.2, we know that

$$\hat{L}(k+h) \triangleq \sum_{l=k+1}^{k+h} \frac{c(l)}{c(k+h)} [L(l) + L'(l)]$$

is a valid Laplacian matrix corresponding to a connected undirected graph. Moreover, given an invertible matrix T with the first column equal to $\mathbf{1}_n$, we have

$$T^{-1} \hat{L}(k+h) T \triangleq \text{diag}\{0, \hat{L}_{(1)}(k+h)\}$$

where $\hat{L}_{(1)}(k+h)$, according to Assumption 8.2 and Lemma A.1, satisfies that $\hat{L}_{(1)}(k+h) > 0$. According to (8.65), we can write

$$\Pi_{k+h, k+1}^{T^{-1}AT} = T^{-1} \Pi_{k+h, k+1}^A T = I - c(k+h) \text{diag}\{0, \hat{L}_{(1)}(k+h)\} + \Delta_\Pi(k+h), \quad (8.66)$$

where $\Delta_\Pi(k+h)$ satisfies that $\|\Delta_\Pi(k+h)\|_2 \leq C_1 \rho_{\bar{d}+h}(k+h)$. According to part 1) of Lemma 8.3, (8.66) implies that $\Pi_{k+lh, k+1}^{T^{-1}AT}$ converge to a matrix with the last $n-1$ rows equal to 0 as $l \rightarrow \infty$. Furthermore,

$$\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^A = T \left(\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^{T^{-1}AT} T^{-1} \right) = \mathbf{1}_n r(k_1),$$

which is an identical row matrix, where $r(k_1) = e_{n,1} \lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^{T^{-1}AT} T^{-1}$. The rest of the proof is to apply the same analysis approach as that in the proof of Theorem 8.1 and Theorem 8.3, which is omitted here. $\square$

Remark 8.4. *Given the distributed protocol (8.61), the consensus is analyzed under the standard Assumption 8.2. We assume that the quotient graph is balanced to guarantee that Π_{k_2, k_1}^A converges to an identical row matrix as $k_2 \rightarrow \infty$. Due to the*

presence of the consensus gain function $c(k)$, the convergence property of Π_{k_2, k_1}^A is difficult to analyze. The traditional approach in [127], [103] cannot be applied since the coefficient of ergodicity is too conservative to be used to prove the result. By using algebraic techniques, we are able to establish that $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^A$ has identical rows under (8.62) in addition to the same conditions as in the fixed topology case.

Remark 8.5. In [65], the mean square consensus for MASs with communication noises and dynamically switching topologies is established under the condition that $c(k)$ monotonically converges to 0 in addition to (8.27), which, however, is more restrictive than the conditions in Theorem 8.5. Theorem 8.5 also proved that when transmission delays happen, the MAS can still reach strong consensus under a weaker condition than that in [65].

8.2.3 Randomly Switching Topologies

We shall consider the case where the communication topology is randomly changing. This problem is strongly motivated by random link failures or random packet dropouts [25], [31], and the random gossip algorithm [30], [5]. The topology to be considered satisfies the following assumption:

Assumption 8.3. The weighting functions $a_{i,j}(k)$, $i, j \in \mathcal{V}$ are randomly distributed over $\mathcal{C} \cup \{0\}$ and $\forall i, j \in \mathcal{V}$, $E(a_{i,j}(k)) = E(a_{j,i}(k))$ is a constant. Moreover, the weighting functions are temporally independent and independent of communication noises $w_{i,j}(k)$.

For dealing with the random link failure problem, the assumption $E(a_{i,j}(k)) = E(a_{j,i}(k))$ is reasonable. It is also a standard assumption for the random gossip algorithm. It is noted that the Laplacian matrix $L(k) = \sum_{l=0}^{\bar{d}} L_l(k)$ of the quotient graph is a random matrix. Define the mean Laplacian matrix as

$$\bar{L} = E(L(k)). \quad (8.67)$$

It is clear that $\bar{L}$ is a valid Laplacian matrix. The graph $\bar{\mathcal{G}}$ corresponding to Laplacian matrix $\bar{L}$ is called mean graph. It is noted that mean graph $\bar{\mathcal{G}}$ is an undirected graph due to that $\bar{L}$ is symmetric. Before giving the result, the following lemma is introduced.

Lemma 8.5. *For any two square matrices M_1 and M_2 with the same dimension and a constant number $\varepsilon > 0$, there is*

$$M_1' M_2 + M_2' M_1 \leq \varepsilon M_1' M_1 + \varepsilon^{-1} M_2' M_2.$$

Proof. The proof is straightforward by considering the following inequality

$$(M_1 - \varepsilon M_2)'(M_1 - \varepsilon M_2) = M_1' M_1 + \varepsilon^2 M_2' M_2 - \varepsilon M_1' M_2 - \varepsilon M_2' M_1 \geq 0.$$

□

Then we have the following result.

Theorem 8.6. *Consider the MAS (8.1)-(8.2) with randomly switching topologies. Given distributed protocol (8.61), under Assumptions 8.1 and 8.3, both strong consensus and mean square consensus are achieved if the mean graph is connected and $c(k)$ satisfies (8.27).*

Proof. We write the system into the form of (8.63)-(8.64). Since all the weighting functions are distributed over $\mathcal{C} \cup \{0\}$, $L(k)$ is uniformly bounded. The system (8.63)-(8.64) can be further written into (8.9) with initial time $k_I = k_0 + \bar{d}$ and initial conditions of $A(k)$ satisfying (8.18), where k_0 is as defined in (8.17) with $\|L\|_\infty$ replaced by $\sup_{k \in \mathbb{Z}^+} \|L(k)\|_\infty$. According to Lemma 8.1, $A(k)$ and $B(k)$ satisfy (8.15) with C as defined in (8.17). Next, we shall investigate the convergence property of Π_{k_2, k_1}^A . Construct the following system:

$$P(k+1) = A(k)P(k). \tag{8.68}$$

Define $\tilde{L}(k) = L(k) - \bar{L}$, where $\bar{L}$ is as defined in (8.67), then the system (8.68) can be written as

$$P(k+1) = (I - c(k)\bar{L} - c(k)\tilde{L}(k) + \Delta_A(k))P(k).$$

It is noted that $E(\tilde{L}(k)) = 0$. Denote $V(k) = P'(k)\bar{L}P(k)$. By simple calculation, we arrive at the following inequality

$$\begin{aligned} E(V(k+1)|P(k)) &= V(k) - 2c(k)P'(k)\bar{L}^2P(k) + \underbrace{2P'(k)\bar{L}E(\Delta_A(k)|P(k))P(k)}_1 \\ &\quad + \underbrace{c^2(k)P'(k)E(\tilde{L}(k)\bar{L}\tilde{L}(k))P(k)}_2 + \underbrace{c^2(k)P'(k)\bar{L}^3P(k)}_3 \\ &\quad - \underbrace{2c(k)P'(k)\bar{L}^2E(\Delta_A(k)|P(k))P(k)}_4 \\ &\quad + \underbrace{P'(k)E(\Delta'_A(k)\bar{L}\Delta_A(k)|P(k))P(k)}_5 \\ &\quad - \underbrace{2c(k)P(k)E(\tilde{L}(k)\bar{L}\Delta_A(k)|P(k))P(k)}_6. \end{aligned}$$

By introducing $\delta_P(k) = \bar{L}^{1/2}P(k)$, it is noted that $\mathbf{1}'_n\delta_P(k) = 0$. Then we have

$$P'(k)\bar{L}^2P(k) = \delta'_P(k)\bar{L}\delta_P(k) \geq \lambda_2(\bar{L})\delta'_P(k)\delta_P(k) = \lambda_2(\bar{L})P'(k)\bar{L}P(k). \quad (8.69)$$

where $\lambda_2(\bar{L})$ is the second smallest eigenvalue of $\bar{L}$. The inequality in (8.69) is due to Lemma A.1. Now we shall show that there exists a constant $\bar{C}$ such that the last 6 terms in the above inequality is bounded by $\rho_{\bar{d}}(k)\bar{C}V(k)$. Consider term 1, according to Lemma 8.5, we have

$$\begin{aligned} &2P'(k)\bar{L}E(\Delta_A(k)|P(k))P(k) \\ &= P'(k)\left[\bar{L}^{1/2}E(\bar{L}^{1/2}\Delta_A(k)|P(k)) + E(\Delta'_A(k)\bar{L}^{1/2}|P(k))\bar{L}^{1/2}\right]P(k) \\ &\leq \rho_{\bar{d}}^{-1}(k)P'(k)E(\Delta'_A(k)|P(k))\bar{L}E(\Delta_A(k)|P(k))P(k) + \rho_{\bar{d}}(k)P'(k)\bar{L}P(k) \\ &\leq \rho_{\bar{d}}(k)(1 + C^2)V(k). \end{aligned} \quad (8.70)$$

The last inequality is due to (8.18). Since $a_{i,j}(k) \in \mathcal{C} \cup \{0\}$, there must exist a constant C_L such that $\sup_{k \geq 0} \|\tilde{L}(k)\|_2 \leq C_L$. Note that $c^2(k) \leq \rho_{\bar{d}}(k)$, then for term 2 we have

$$c^2(k)P'(k)E(\tilde{L}(k)\bar{L}\tilde{L}(k))P(k) \leq \rho_{\bar{d}}(k)C_L^2V(k). \quad (8.71)$$

According to Lemma A.1 and (8.69), term 3 satisfies

$$c^2(k)P'(k)\bar{L}^3P(k) \leq \rho_{\bar{d}}(k)\lambda_2^2(\bar{L})V(k). \quad (8.72)$$

Similarly to (8.70), (8.71) and (8.72), we arrive at the inequality for term 4, 5, 6 as follows

$$\begin{aligned} & 2c(k)P'(k)\bar{L}^2E(\Delta_A(k)|P(k))P(k) \\ &= c(k)P'(k) \left[\bar{L}^{3/2}E(\bar{L}^{1/2}\Delta_A(k)|P(k)) + E(\Delta'_A(k)\bar{L}^{1/2}|P(k))\bar{L}^{3/2} \right] P(k) \\ &\leq K\rho_{\bar{d}}(k)P'(k)\bar{L}^3P(k) + K\rho_{\bar{d}}^{-1}(k)P'(k)E(\Delta'_A(k)|P(k))\bar{L}E(\Delta_A(k)|P(k))P(k) \\ &\leq \rho_{\bar{d}}(k)K \left(\lambda_2^2(\bar{L}) + C^2 \right) V(k), \end{aligned} \quad (8.73)$$

$$P'(k)E(\Delta'_A(k)\bar{L}\Delta_A(k)|P(k))P(k) \leq C^2\rho_{\bar{d}}^2(k)V(k) \leq \rho_{\bar{d}}(k)K^2C^2V(k), \quad (8.74)$$

$$\begin{aligned} & 2c(k)P'(k)E(\tilde{L}(k)\bar{L}\Delta_A(k)|P(k))P(k) \\ &= c(k)P'(k)E \left((\tilde{L}(k)\bar{L}^{1/2})(\bar{L}^{1/2}\Delta_A(k)) + (\Delta'_A(k)\bar{L}^{1/2})(\bar{L}^{1/2}\tilde{L}(k)) \right) P(k) \\ &\leq KP'(k) \left[\rho_{\bar{d}}(k)E(\tilde{L}(k)\bar{L}\tilde{L}(k)) + \rho_{\bar{d}}^{-1}(k)E(\Delta'_A(k)\bar{L}\Delta_A(k)|P(k)) \right] P(k) \\ &\leq \rho_{\bar{d}}(k)K(C_L^2 + C^2)V(k). \end{aligned} \quad (8.75)$$

By defining $\bar{C} = 1 + C^2 + C_L^2 + \lambda_2^2 + K(\lambda_2^2 + 2C^2 + C_L^2)$, then we have $0 < \bar{C} < +\infty$. Furthermore,

$$E(V(k+1)|P(k)) \leq (1 - 2\lambda_2(\bar{L})c(k) + \bar{C}\rho_{\bar{d}}(k))V(k). \quad (8.76)$$

Since the mean graph is connected according to Assumption 8.3, we know that

$\lambda_2(\bar{L}) > 0$ by Lemma A.1. Hence, (8.76) implies that

$$\lim_{k \rightarrow \infty} V(k) = \lim_{k \rightarrow \infty} P'(k) \bar{L} P(k) = 0$$

almost surely [66], [53]. According to the definition of V we have $\lim_{k \rightarrow \infty} P(k) = \mathbf{1}_n y^*$, where y^* is a constant. In other words, Π_{k_2, k_1}^A converges to an identical row matrix as $k_2 \rightarrow \infty$ almost surely. Along the same line as the proof of Theorem 8.1, we arrive at that $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^\Phi = [\mathbf{1}'_n \mathbf{0}'_{n(\bar{d}+1)}]' [r_1(k_1) \ r_2(k_1)]$ almost surely. The rest of the proof is to apply the same analysis approach as that in the proof of Theorem 8.1 and Theorem 8.3, which is omitted here. $\square$

Remark 8.6. *The key to our analysis of the consensus problem under randomly switching topologies is still the convergence property of Π_{k_2, k_1}^A . In this case, A becomes a random matrix. Based on the technique in [53], we proved that Π_{k_2, k_1}^A converges almost surely to a matrix with identical rows as $k_2 \rightarrow \infty$, which allowed us to extend the results obtained in Section 8.2 to the randomly switching topologies case.*

Remark 8.7. *It is easy to show that under dynamical switching topology and random switching topology, the consensus condition (8.27) on $c(k)$ can be relaxed to (8.25) if the communication topology is further with leader-follower structure.*

8.3 Numerical Examples

Let $\mathcal{V} = \{1, 2, 3, 4\}$, $\mathcal{E}(\mathcal{G}_Q) = \{(1, 2), (2, 3), (3, 4)\}$ and the quotient graph be shown in Fig. 8.1. All the measurement noises are uniformly distributed over $[-1, 1]$. The upper bound of delay is $\bar{d} = 2$. The initial conditions are:

$$\begin{aligned} x_1(1) &= 1, \ x_2(1) = 2.1, \ x_3(1) = 1, \quad x_4(1) = 3, \\ x_1(2) &= 1, \ x_2(2) = -1, \ x_3(2) = -0.5, \ x_4(2) = 1.5, \\ x_1(3) &= 1, \ x_2(3) = -1, \ x_3(3) = -0.5, \ x_4(3) = 0. \end{aligned}$$

Figure 8.1: Quotient graph of the MAS.

We assume the delays to be stochastic with the probability distribution given as follows:

$$\begin{aligned}
 P(d_{1,2} = 0) &= 0.2, \quad P(d_{1,2} = 1) = 0.4, \quad P(d_{1,2} = 2) = 0.4, \\
 P(d_{2,3} = 0) &= 0.5, \quad P(d_{2,3} = 1) = 0.3, \quad P(d_{2,3} = 2) = 0.2, \\
 P(d_{3,1} = 0) &= 0.1, \quad P(d_{3,1} = 1) = 0.4, \quad P(d_{3,1} = 2) = 0.5, \\
 P(d_{1,4} = 0) &= 0.3, \quad P(d_{1,4} = 1) = 0.3, \quad P(d_{1,4} = 2) = 0.4.
 \end{aligned}$$

The consensus protocol (8.5) is applied with $a_{i,j} = 1$, $t_{i,j} = 0$ for all $i \in \mathcal{V}$ and $j \in \mathcal{N}_i$. Note that the topology is with leader-follower structure. When we set $c(k) = 0.5$, the effect of the noises cannot be attenuated, which can be seen from the state trajectories in Fig. 8.2. If we set $c(k) = (k + \bar{d})^{-0.5}$, which satisfies (8.25), the states will reach a consensus as shown in Fig. 8.3.

Next, we consider the case that the communication graph is an undirected version of the graph in Fig. 8.1, i.e. the edges in Fig. 8.1 are bidirectional. The topology becomes a non-leader-follower structure. If we still choose $c(k) = (k + \bar{d})^{-0.5}$, according to Theorem 8.1 and 8.3, the consensus fails since $\sum_{k=0}^{\infty} c^2(k) = +\infty$. The states are shown in Fig. 8.4. We can see that the states cannot converge. By choosing $c(k) = (k + \bar{d})^{-1}$, the condition (8.27) is satisfied. The corresponding state responses in Fig. 8.5 demonstrate the consensus. When the communication graph is time-varying according to the following rules

$$\begin{aligned}
 \mathcal{E}(\mathcal{G}_Q(k)) &= \{(1, 2), (3, 4)\}, \quad \text{when } k \text{ is even,} \\
 \mathcal{E}(\mathcal{G}_Q(k)) &= \{(2, 3)\}, \quad \text{when } k \text{ is odd,}
 \end{aligned}$$

Figure 8.2: Leader-follower topology with $c(k) = 0.5$

Figure 8.3: Leader-follower topology with $c(k) = 1/\sqrt{k+d}$

then by Theorem 8.5, the states can reach a consensus under the protocol (8.61). In the case where the edge $(2, 3)$ of the fixed graph is randomly broken in the following way

$$P((2, 3) \in \mathcal{E}) = P((2, 3) \notin \mathcal{E}) = 0.5,$$

according to Theorem 8.6 the states also converge to a random variable under the same protocol (8.61). The state trajectories under the dynamically switching topologies and randomly switching topologies are shown in Fig. 8.6 and Fig. 8.7, respectively.

Figure 8.4: Fixed undirected topology with $c(k) = 1/\sqrt{k+d}$

8.4 Conclusion

The consensus problem of MASs with communication noises and transmission delays has been considered. Due to the existence of delays and noises, the problem becomes rather challenging. Traditional approaches such as the stochastic Lyapunov analysis, supermartingale convergence theory and double array analysis cannot be directly applied to analyze the convergence property and consensusability for the class of systems. A novel technique is introduced for the first time to represent the system in an alternative form which leads to a feasible way for analysis. Necessary and sufficient conditions for strong consensus and mean square consensus under fixed topology have been provided. The consensus problem under dynamically switching

Figure 8.5: Fixed undirected topology with $c(k) = 1/(k + \bar{d})$

Figure 8.6: Dynamically switching topologies with $c(k) = 1/(k + \bar{d})$

topologies and randomly switching topologies has also been investigated. Some numerical examples have been given to demonstrate the effectiveness of the results.

Figure 8.7: Randomly switching topologies with $c(k) = 1/(k + \bar{d})$

8.5 Appendix: Proof of Lemma 8.4

Before proving Lemma 8.3, the following lemma is introduced.

Lemma 8.6. *For any number $\beta_1 \in \mathbb{R}$ and $0 < \beta_2 \leq \alpha < 1$, the following inequality holds*

$$1 - \beta_1 \leq e^{-\beta_1}, \quad (8.77)$$

$$1 - \beta_2 \geq e^{-\tau\beta_2}, \quad (8.78)$$

where

$$\tau = \frac{-\ln(1 - \alpha)}{\alpha}. \quad (8.79)$$

Proof. Define a function $f_1(\beta) = 1 - \beta - e^{-\beta}$. It is easy to check that $f_1(0) = 0$ and $\dot{f}_1(\beta) < 0, \forall \beta > 0; \dot{f}_1(\beta) > 0, \forall \beta < 0$. Thus (8.77) holds.

Next, define $f_2(\beta) = 1 - \beta - e^{-\tau\beta}$. Then we have

$$f_2(0) = f_2(\alpha) = 0. \quad (8.80)$$

Moreover, $\ddot{f}_2(\beta) = -\tau^2 e^{-\tau\beta} < 0$, which together with (8.80) leads to (8.78). $\square$

Proof of Lemma 8.3 1). We denote the eigenvalues of L by λ_i , $i = 1, \dots, n$. Without loss of generality, let the eigenvalues of L satisfy that $0 < \operatorname{Re}(\lambda_i) \leq \operatorname{Re}(\lambda_j)$ when $i < j$. For simplicity, we assume that L can be expressed as $L = T_1 \Lambda T_1^{-1}$, where $\Lambda = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\} > 0$. It is easy to extend the proof to the general case in which Λ is of a Jordan canonical form. We just consider the case that all the eigenvalues are real numbers. The case that there exist complex eigenvalues will not increase any difficulty for analysis but make notation complicated.

$M(k)$ can be written as $M(k) = T_1 M_1(k) T_1^{-1}$, where $M_1(k) = I - c(k)\Lambda + \Delta_{M_1}(k)$, $\Delta_{M_1}(k) = T_1^{-1}(M(k) - I + c(k)L)T_1$. By denoting $C_{T_1} = \|T_1\|_2 \|T_1^{-1}\|_2$, we can see that $\|\Delta_{M_1}(k)\|_2 \leq C_1 C_{T_1} \rho_m(k)$. By choosing k_I such that $\forall k \geq k_I$, $c(k)\Lambda < I$, we have

$$\|M_1(k)\|_2 \leq 1 - c(k)\lambda_1 + C_1 C_{T_1} \rho_m(k), \quad \forall k \geq k_I.$$

Hence, $\forall k_1 \geq k_I$,

$$\|\Pi_{k_2, k_1}^M\|_2 = \|T_1 \Pi_{k_2, k_1}^{M_1} T_1^{-1}\|_2 \leq C_T \|\Pi_{k_2, k_1}^{M_1}\|_2 \leq C_{T_1} \prod_{k=k_1}^{k_2} (1 - c(k)\lambda_1 + C_1 C_{T_1} \rho_m(k)).$$

By letting $C_2 = C_{T_1}$, $C_3 = \lambda_1$, $C_4 = C_1 C_{T_1}$, (8.26) is obtained.

Since $\lim_{k \rightarrow \infty} c(k) = 0$, by considering (8.14), there exists k_3 such that $\forall k \geq k_3$, $C_4 \rho_m(k) < c(k)C_3/2$. Therefore,

$$\begin{aligned} \lim_{k_2 \rightarrow \infty} \|\Pi_{k_2, k_1}^M\|_2 &\leq C_2 \prod_{k=k_1}^{\infty} (1 - C_3 c(k) + C_4 \rho_m(k)) \\ &= C_2 \prod_{k=k_1}^{k_3-1} (1 - C_3 c(k) + C_4 \rho_m(k)) \prod_{k=k_3}^{\infty} (1 - C_3 c(k) + C_4 \rho_m(k)) \\ &\leq C_2 e^{-\frac{C_3}{2} \sum_{k=k_3}^{\infty} c(k)} \prod_{k=k_1}^{k_3-1} (1 - C_3 c(k) + C_4 \rho_m(k)) \\ &= 0. \end{aligned} \tag{8.81}$$

The inequality in (8.81) comes from Lemma (8.6). (8.81) implies that Π_{k_2, k_1}^M converges to 0 as $k_2 \rightarrow \infty$.

2). Since $\lim_{k \rightarrow \infty} c(k) = 0$, we can choose k_I large enough such that $\forall k \geq k_I$, $1 - c(k) \|L_{(1)}\|_2 > 0$. Therefore, $\forall k_1 \geq k_I$,

$$\lim_{k_2 \rightarrow \infty} \|\Pi_{k_2, k_1}^M\|_2 \leq \Pi_{k=k_1}^\infty (1 + C_1 \rho_m(k)) < e^{C_1 \sum_{k=k_1}^\infty \rho_m(k)} < +\infty. \quad (8.82)$$

The last inequality is due to that $\sum_{k=k_1}^\infty \rho_m(k) \leq (\bar{d} + 1) \sum_{k=k_1}^\infty c^2(k) < +\infty$. Then we construct the following system

$$P(k+1) = M(k)P(k), \quad (8.83)$$

where $P(k) = \text{col}\{p_1(k), p_2(k), \dots, p_n(k)\}$ is a column vector. (8.82) implies that $P(k)$ is uniformly bounded for any bounded initial condition. By denoting $P_2(k) = \text{col}\{p_2(k), \dots, p_n(k)\}$, we have

$$P_2(k+1) = M_{(1)}(k)P_2(k) + \delta(k)p_1(k), \quad (8.84)$$

where $\delta(k) = \text{col}\{\Delta_M^{2,1}(k), \Delta_M^{3,1}(k), \dots, \Delta_M^{n,1}(k)\}$,

$$M_{(1)}(k) = I - c(k)L_{(1)} + \Delta_{M_{(1)}}(k), \quad (8.85)$$

and $\Delta_{M_{(1)}}(k)$ is obtained from $\Delta_M(k)$ by deleting the first column and first row. Since all the eigenvalues of $L_{(1)}$ lie on the open right-hand plane, according to 1), we know that $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^{M_{(1)}} = 0$ for all $k_1 \geq 0$. Given $P(k_1)$, denote $R_1 = \sup_{k_2 > k \geq k_1} \|\Pi_{k_2, k+1}^{M_{(1)}} p_1(k)\|_2$. It is obvious that $\|\delta(k)\|_2 \leq C_1 \rho_m(k)$. Hence, for any given arbitrarily small number $\varepsilon > 0$, there exists a number $N(\varepsilon) \in \mathbb{Z}^+$ such that

$$\sum_{k=N(\varepsilon)}^\infty \|\delta(k)\|_2 < \frac{\varepsilon}{R_1}. \quad (8.86)$$

Therefore,

$$\begin{aligned}
\lim_{k_2 \rightarrow \infty} \|P_2(k_2 + 1)\|_2 &= \lim_{k_2 \rightarrow \infty} \left\| \Pi_{k_2, k_1}^{M(1)} P_2(k_1) + \sum_{l=k_1}^{k_2} \Pi_{k_2, l+1}^{M(1)} p_1(l) \delta(l) \right\|_2 \\
&\leq \lim_{k_2 \rightarrow \infty} \left\| \Pi_{k_2, k_1}^{M(1)} P_2(k_1) + \sum_{l=k_1}^{N(\varepsilon)-1} \Pi_{k_2, l+1}^{M(1)} p_1(l) \delta(l) \right\|_2 \\
&\quad + \lim_{k_2 \rightarrow \infty} \left\| \sum_{l=N(\varepsilon)}^{k_2} \Pi_{k_2, l+1}^{M(1)} p_1(l) \delta(l) \right\|_2 \\
&\leq R_1 \sum_{l=N(\varepsilon)}^{\infty} \|\delta(l)\|_2 \\
&< \varepsilon.
\end{aligned}$$

The arbitrariness of ε implies that $\lim_{k_2 \rightarrow \infty} P_2(k_2) = 0$ for any bounded initial condition $P(k_1)$. According to (8.83), we conclude that the last $n - 1$ rows of $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^M$ are 0.

Finally, we shall show that, when $\sum_{k=0}^{\infty} c(k) < +\infty$, $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^M$ cannot converge to a matrix with the last $n - 1$ rows equal to 0. Denote

$$R_2 \triangleq \sup_{k_2 \geq k_1 \geq k_I} \|\Pi_{k_2, k_1}^M\|_2, \quad (8.87)$$

then $\|P(k_2)\|_2 \leq R_2 \|P(k_1)\|_2$, $\forall k_2 \geq k_1 \geq k_I$.

Next, we will prove that there exists k'_I such that $\inf_{k_2 \geq k_1 \geq k'_I} \|\Pi_{k_2, k_1}^{M(1)}\|_2$ is bounded away from 0. Actually, we can define

$$k'_I = k_I \vee s \left(\frac{\alpha}{2\|L_{(1)}\|_2} \wedge \sqrt{\frac{\alpha}{2C_1}} \right),$$

where $s(\cdot)$ is as defined in (8.16). Clearly,

$$0 \leq c(k) \|L_{(1)}\|_2 + C_1 \rho_m(k) \leq \alpha, \quad \forall k \geq k'_I. \quad (8.88)$$

According to the definition of $M_{(1)}(k)$ in (8.85), we know that $\|M_{(1)}(k)\|_2 \leq$

$C_1\rho_m(k)$, and

$$\left\| \left(\Pi_{k_2, k_1}^{M_{(1)}} \right)^{-1} \right\|_2 \leq \prod_{k=k_1}^{k_2} \left\| M_{(1)}^{-1}(k) \right\|_2 \leq \prod_{k=k_1}^{k_2} (1 - c(k)\|L_{(1)}\|_2 - C_1\rho_m(k))^{-1}. \quad (8.89)$$

The last inequality follows Lemma 8.2. By considering (8.88) and applying Lemma 8.6, we have

$$1 - c(k)\|L_{(1)}\|_2 - C_1\rho_m(k) \geq e^{-\tau[c(k)\|L_{(1)}\|_2 + C_1\rho_m(k)]}, \quad \forall k \geq k'_I,$$

where τ is as defined in (8.79), which together with (8.89) leads to that

$$\begin{aligned} \sup_{k_2 \geq k_1 \geq k'_I} \left\| \left(\Pi_{k_2, k_1}^{M_{(1)}} \right)^{-1} \right\|_2^{-1} &\geq \sup_{k_2 \geq k_1 \geq k'_I} \prod_{k=k_1}^{k_2} (1 - c(k)\|L_{(1)}\|_2 - C_1\rho_m(k)) \\ &\geq e^{-\tau \sum_{k=k'_I}^{\infty} [c(k)\|L_{(1)}\|_2 + C_1\rho_m(k)]} \triangleq \underline{U}. \end{aligned} \quad (8.90)$$

Due to that $\sum_{k=0}^{\infty} c(k) < +\infty$, $\sum_{k=0}^{\infty} \rho_m(k) < +\infty$, we know that $\underline{U} > 0$. On the other hand,

$$\left\| \left(\Pi_{k_2, k_1}^{M_{(1)}} \right)^{-1} \right\|_2 \left\| \Pi_{k_2, k_1}^{M_{(1)}} \right\|_2 \geq 1.$$

Thus, we have

$$\inf_{k_2 \geq k_1 \geq k'_I} \left\| \Pi_{k_2, k_1}^{M_{(1)}} \right\|_2 \geq \sup_{k_2 \geq k_1 \geq k'_I} \left\| \left(\Pi_{k_2, k_1}^{M_{(1)}} \right)^{-1} \right\|_2^{-1} \geq \underline{U}.$$

Hence, $\Pi_{k_2, k_1}^{M_{(1)}}$ cannot converge to 0 as $k_2 \rightarrow \infty$. There must exist a vector $v(k_1)$ such that

$$\lim_{k_2 \rightarrow \infty} \left\| \Pi_{k_2, k_1}^{M_{(1)}} v(k_1) \right\|_2 \geq \underline{U}, \quad \|v(k_1)\|_2 = 1. \quad (8.91)$$

Also, we can prove that $\Pi_{k_2, k_1}^{M_{(1)}}$ is bounded as $k_2 \rightarrow \infty$. In fact,

$$\sup_{k_2 \geq k_1 \geq k'_I} \left\| \Pi_{k_2, k_1}^{M_{(1)}} \right\|_2 \leq \sup_{k_2 \geq k_1 \geq k'_I} \prod_{k=k_1}^{k_2} (1 + c(k)\|L_{(1)}\|_2 + C_1\rho_m(k))$$

$$\leq e^{\sum_{k=k'_I}^{\infty} [c(k)\|L_{(1)}\|_2 + C_1\rho_m(k)]} \triangleq \bar{U}, \quad (8.92)$$

Obviously, $\bar{U} = \underline{U}^{-\frac{1}{\tau}} < +\infty$.

Next, since $\|\delta(k)\|_2 \leq C_1\rho_m(k)$, there must exist a number $N \in \mathbb{Z}^+$ such that

$$\sum_{k=N}^{\infty} \|\delta(k)\|_2 < \frac{\underline{U}}{\bar{U}R_2}, \quad (8.93)$$

where R_2 , $\underline{U}$ and $\bar{U}$ are as defined in (8.87), (8.90) and (8.92). Then we consider the system (8.83) starting from time instant $k_1 \triangleq k'_I \vee N$. By choosing initial value

$$P(k_1) = [0 \ v'(k_1)]', \quad (8.94)$$

with v defined in (8.91) and considering (8.84) we have

$$\begin{aligned} \lim_{k \rightarrow \infty} \|P_2(k)\|_2 &= \lim_{k \rightarrow \infty} \left\| \Pi_{k-1, k_1}^{M_{(1)}} P_2(k_1) + \sum_{l=k_1}^{k-1} \Pi_{k-1, l+1}^{M_{(1)}} \delta(l) p_1(l) \right\|_2 \\ &\geq \lim_{k \rightarrow \infty} \left\| \Pi_{k-1, k_1}^{M_{(1)}} v(k_1) \right\|_2 - \lim_{k \rightarrow \infty} \sum_{l=k_1}^{k-1} \left\| \Pi_{k-1, l+1}^{M_{(1)}} \delta(l) p_1(l) \right\|_2 \\ &\geq \underline{U} - \bar{U} R_2 \|P(k_1)\|_2 \sum_{l=k_1}^{\infty} \|\delta(l)\|_2 \\ &> 0. \end{aligned} \quad (8.95)$$

The second inequality is due to (8.91) and $\|p_1(k)\|_2 \leq \|P(k)\|_2 \leq R_4 \|P(k_1)\|_2$. The last inequality comes from (8.93) and (8.94). (8.95) implies that Π_{k_2, k_1}^M cannot converge to a matrix with the last $n-1$ rows equal to 0 as $k_2 \rightarrow \infty$.

3). Consider $M_{(1)}(k)$ as defined in (8.85) with $L_{(1)}$ becoming a time-varying matrix. It is clear that $\|\Delta_{M_{(1)}}(k)\|_2 \leq C_1\rho_m(k)$. Since the eigenvalues of $L(k) + L'(k)$ are uniformly positive, we denote

$$\underline{\lambda} = \inf_{k \geq 0} \lambda(L_{(1)}(k) + L'_{(1)}(k)) > 0, \quad \bar{\lambda} = \sup_{k \geq 0} \lambda(L_{(1)}(k) + L'_{(1)}(k)) < +\infty.$$

By setting $k_I = m \vee s(1/\bar{\lambda})$ with $s(\cdot)$ as defined in (8.16), we have $\forall k \geq k_I$,

$$\begin{aligned}
\|M_{(1)}(k)M'_{(1)}(k)\|_2 &= \left\| [I - c(k)L_{(1)}(k) + \Delta_{M_{(1)}}(k)][I - c(k)L_{(1)}(k) + \Delta_{M_{(1)}}(k)]' \right\|_2 \\
&\leq \|I - c(k)[L_{(1)}(k) + L'_{(1)}(k)]\|_2 + \|\Delta_{M_{(1)}}(k)\|_2 + \|\Delta'_{M_{(1)}}(k)\|_2 \\
&\quad + \|c(k)L_{(1)}(k)\Delta_{M'_{(1)}}(k)\|_2 + \|c(k)\Delta_{M_{(1)}}(k)L'_{(1)}(k)\|_2 \\
&\quad + \|c^2(k)L_{(1)}(k)L'_{(1)}(k)\|_2 + \|\Delta_{M_{(1)}}(k)\Delta'_{M_{(1)}}(k)\|_2 \\
&\leq 1 - \underline{\lambda}c(k) + 2C_1\rho_m(k) + 2\bar{c}R_3C_1\rho_m(k) + R_3^2\rho_m(k) \\
&\quad + \bar{c}^2C_1^2\rho_m(k).
\end{aligned} \tag{8.96}$$

where $\bar{c} = \sup_{k \geq 0} c(k)$, $R_3 = \sup_{k \in \mathbb{Z}^+} \|L(k)\|_2$. Furthermore,

$$\begin{aligned}
\left\| \left(\Pi_{k_2, k_1}^{M_{(1)}} \right)' \Pi_{k_2, k_1}^{M_{(1)}} \right\|_2 &\leq \prod_{k=k_1}^{k_2} \|M_{(1)}(k)M'_{(1)}(k)\|_2 \\
&\leq \prod_{k=k_1}^{k_2} [1 - \underline{\lambda}c(k) + (2C_1 + 2\bar{c}R_3C_1 + R_3^2 + \bar{c}^2C_1^2)\rho_m(k)].
\end{aligned}$$

The last inequality comes from (8.96). By applying same argument as that in 1), we can prove that $\Pi_{k_2, k_1}^{M_{(1)}}$ converges to 0 as $k_2 \rightarrow \infty$. Further, by using the same procedure as that in the proof of 2), we can prove that the last $n - 1$ rows of $\lim_{k_2 \rightarrow \infty} \Pi_{k_2, k_1}^M$ are 0. $\square$

Chapter 9

Distributed Consensus for MAS with Communication Delays and Limited Data Rate

It can be seen that most of the existing works deal with the consensus problem with communication delays and limited data rate separately. Note that time delays and data-rate constraint co-exist in real multi-agent networks. However, for networks with both communication delays and data-rate constraint, the average consensus problem can not be solved by the existing methods. Moreover, for the finite-level quantization, quantization errors can not be regarded as additive white noises, so the method in Chapter 8 is not applicable. In this chapter, we shall consider the average consensus problem for undirected networks with both communication delays and data-rate constraint.

We adopt a dynamic encoding and decoding scheme with finite-level quantizers. We design a distributed protocol with error compensation. Note that, the time delays are to be considered in the encoder-decoder and the protocol design. If we write the whole system into an augmented system, it can be seen that the system is time-varying and the system contains multiple delays in the states. The existing results cannot be applied to analyze the convergence of the augmented system. To

overcome this difficulty, an auxiliary system is introduced to represent the system in two coupled parts, a persistent part which converges to a delay free version of the original systems and a delay part which contains all the effects of the delays. We can get a stability condition which is given in terms of the network topology, the upper bound of the time delay, the number of quantization levels and the scaling function to be designed for coding. We show that if the network is connected and the time delays are bounded, then for any given finite-level uniform quantizer, the consensus gain and the scaling function can be chosen properly such that the average consensus can be achieved asymptotically. In particular, if the control parameters are properly chosen, the average consensus can be guaranteed with merely one bit information exchange between agents at each time step.

9.1 Problem Statement

In this chapter, we consider N agents with the following dynamics:

$$x_i(k+1) = x_i(k) + u_i(k), \quad i = 1, 2, \dots, N. \quad (9.1)$$

The communication topology among agents is modeled as a graph $\mathcal{G}$.

In this chapter, we assume that transmission channels are noiseless digital channels without packet dropouts. The time delay of the transmission channel from i to j is denoted by $d_{i,j}$, where $d_{i,j}$ is a nonnegative integer. The communication between each pair of neighbors are as follows: at each time instant k , the agent i encodes its state $x_i(k)$ by an encoder ϕ_i into a symbolic data $\Delta_i(k)$, which will be sent to agent j through the channel. Due to the time delay in the transmission channel from i to j , the symbolic data received by the neighbor node j at time instant k is $\Delta_i(k - d_{j,i}(k))$. Then agent j uses a decoder $\varphi_{i,j}$ to get an estimate of agent i 's state. The whole communication process from agent i to j is shown in Fig. 1.

We say $\{u_1(t), u_2(t), \dots, u_N(t), t = 0, 1, \dots\}$ is a distributed protocol over digital network $\mathcal{G}$, if $u_i(t)$ only depends on the states of itself, the state of encoder ϕ_i and the

Figure 9.1: The transmission channel from i to j .

outputs of decodes $\varphi_{j,i}$, $j \in N_i$. In this chapter, we will consider how to design proper encoders, decoders and a distributed protocol $\{u_1(t), u_2(t), \dots, u_N(t), t = 0, 1, \dots\}$ such that

$$\lim_{k \rightarrow \infty} x_i(k) = \frac{1}{N} \sum_{i=1}^N x_i(0), \quad i = 1, 2, \dots, N.$$

We introduce the following assumptions to communication channels and topology.

Assumption 9.1. *The communication graph $\mathcal{G}$ is a connected undirected graph.*

Assumption 9.2. *The time delays of all channels are bounded and symmetric, i.e. $d_{i,j}(k) = d_{j,i}(k)$, and $\max_{(i,j) \in \mathcal{E}} d_{i,j}(k) \leq d$, where d is a positive integer.*

Assumption 9.3. *Each agent has a memory of its past d states, and data transmitted through a channel is with time stamp.*

Assumption 9.4. *There is a known constant C_x , such that $\max_i |x_i(0)| \leq C_x$.*

Remark 9.1. *We assume that the delays are symmetric in each transmission channel in order to guarantee that the average value of the initial states are preserved at each time step. Meanwhile, we need an upper bound C_x of the initial states to design the protocol. Note that, C_x can be estimated in a distributed way.*

9.2 Protocol Design and Consensus Analysis

9.2.1 Protocol Design

It is noted that the exact consensus cannot be reached by using static quantizers [55]. Therefore, we need to design a dynamic encoder and decoder with finite-level

Figure 9.2: The encoder ϕ_i .

quantization to process the data to be transmitted. The encoder for agent i is given as follows,

$$\begin{cases} \hat{x}_i(k) = g(k-1)\Delta_i(k) + \hat{x}_i(k-1), \hat{x}_i(0) = 0, \\ \Delta_i(k) = q \left[\frac{x_i(k) - \hat{x}_i(k-1)}{g(k-1)} \right], \end{cases} \quad (9.2)$$

where $\hat{x}_i(k)$ is the encoder state, $g(k) > 0$ is the scaling function to be designed, $q[\cdot]$ is a quantizer. Fig. 9.2 shows the diagram of the encoder (9.2). In fact, $\hat{x}_i$ can be viewed as the estimation of the state and the inverse scaling function $\frac{1}{g(k)}$ is to zoom in the estimation error $e_i(k) \triangleq x_i(k+1) - \hat{x}_i(k)$. Then we quantize the zoomed innovation and send the quantized data $\Delta_i(k)$ to transmission channel. As shown above, the quantized data received by the neighbor node j at time instant k is $\Delta_i(k - d_{j,i}(k))$. We propose a decoder for agent j associated with the transmission channel from i to j of the form:

$$\hat{x}_{j,i}(k) = \hat{x}_{j,i}(k-1) + g(k - d_{j,i}(k) - 1)\Delta_i(k - d_{j,i}(k))\chi_{[0,\infty)}(k - d_{j,i}(k)), \hat{x}_{j,i}(0) = 0. \quad (9.3)$$

The diagram of the decoder is given in Fig. 9.3. It is straightforward that $\hat{x}_{j,i}(k) = \hat{x}_i(k - d_{j,i}(k))$.

The scaling function $g(\cdot)$ should satisfy that $\lim_{k \rightarrow \infty} g(k) = 0$. Here, we choose exponential function $g_0\gamma^k$ with $0 < \gamma < 1$ as the scaling function $g(k)$. We consider

Figure 9.3: The decoder $\varphi_{i,j}$.

a finite-level uniform quantizer which is defined as follows:

$$q[x] = \begin{cases} K, & x \geq K - \frac{1}{2}, \\ l, & l - \frac{1}{2} \leq x < l + \frac{1}{2}, \\ & l = 1, 2, \dots, K-1, \\ 0, & -\frac{1}{2} < x < \frac{1}{2}, \\ -q[-x], & x \leq -\frac{1}{2}. \end{cases} \quad (9.4)$$

From (9.4) we can see that the number of quantization levels is $2K + 1$. No information will be sent when the output of quantizer is 0, in which case the communication channel from agent i to j , $j \in \mathcal{N}_i$ is required to be capable of transmitting $\lceil \log_2(2K) \rceil$ bits.

We propose a distributed protocol as

$$u_i(k) = h \sum_{j \in \mathcal{N}_i} a_{i,j} \chi_{[0,\infty)}(k - d_{i,j}(k)) [\hat{x}_{i,j}(k) - \hat{x}_i(k - d_{i,j}(k))], \quad (9.5)$$

where h is the consensus gain function to be designed.

9.2.2 Consensus Analysis

Define the sub-Laplacian matrix $L_l(k)$ as

$$L_l^{i,j}(k) = \begin{cases} -a_{i,j} \delta_{d_{i,j}(k), l} \chi_{[l,\infty)}(k), & i \neq j, \\ \sum_{j \in \mathcal{N}_i} a_{i,j} \delta_{d_{i,j}(k), l} \chi_{[l,\infty)}(k), & i = j, \end{cases}$$

where $L_l(k)$ is the Laplacian matrix of the subgraph ([10]) corresponding to time delay l . It is noted that $L_l(k) = L'_l(k)$ thanks to the symmetry of delays. According to the definition of $L_l(k)$, we have

$$L_l(k)\mathbf{1}_N = 0, \quad \mathbf{1}'_N L_l(k) = 0, \quad l = 0, 1, \dots, d. \quad (9.6)$$

Denote

$$\begin{aligned} X(k) &= \text{col} \{x_1(k), x_2(k), \dots, x_N(k)\}, \quad \hat{X}(k) = \text{col} \{\hat{x}_1(k), \hat{x}_2(k), \dots, \hat{x}_N(k)\}, \\ \tilde{X}(k) &= X(k) - \hat{X}(k), \quad \xi(k) = X(k) - J_N X(k), \end{aligned}$$

where $J_N = \mathbf{1}_N \mathbf{1}'_N / N$. It is noted that $\xi(k)$ stands for the disagreement, i.e. the deviations of the states from the average value at time instant k . The closed-loop system can be written in the following compact form

$$\begin{cases} X(k+1) = X(k) - h \sum_{i=0}^d L_i(k) \hat{X}(k-i), \\ \hat{X}(k+1) = \hat{X}(k) + g(k)q \left[\frac{X(k+1) - \hat{X}(k)}{g(k)} \right], \end{cases} \quad (9.7)$$

where the quantizer $q[\cdot]$ is carried out elementwisely. According to (9.6) and (9.7) we get

$$\frac{1}{N} \sum_{i=1}^N x_i(k+1) = \frac{1}{N} \mathbf{1}'_N X(k+1) = \frac{1}{N} \mathbf{1}'_N X(k) = \frac{1}{N} \sum_{i=1}^N x_i(k), \quad (9.8)$$

which means that the state average is preserved at each step. Thus, the average consensus problem is solved if and only if

$$\lim_{k \rightarrow \infty} \xi(k) = 0. \quad (9.9)$$

In view of (9.6) and (9.7), we can easily see that

$$e(k) \triangleq X(k+1) - \hat{X}(k)$$

$$\begin{aligned}
&= \tilde{X}(k) + h \sum_{i=0}^d L_i(k) [\tilde{X}(k-i) - X(k-i)] \\
&= \tilde{X}(k) + h \sum_{i=0}^d L_i(k) [\tilde{X}(k-i) - \xi(k-i)]. \tag{9.10}
\end{aligned}$$

Note that $(I - J_N)L_i(k) = L_i(k) = L_i(k)(I - J_N)$, which together with (9.7) and (9.10) yields

$$\begin{cases} \xi(k+1) = \xi(k) - h \sum_{i=0}^d L_i(k)\xi(k-i) + h \sum_{i=0}^d L_i(k)\tilde{X}(k-i), & (9.11a) \\ \tilde{X}(k+1) = e(k) - g(k)q \left[\frac{e(k)}{g(k)} \right]. & (9.11b) \end{cases}$$

It is noted that there are multiple delays in system (9.11a) and the matrices $L_i(k)$, $i = 0, \dots, d$ are time-varying, which poses challenges for consensus analysis. We shall provide a way to choose the consensus gain h , the parameters g_0 and γ of the scaling function and quantization level $2K + 1$ such that (9.9) holds. By looking at (9.11a)-(9.11b), it is easy to see that (9.9) holds when $\lim_{k \rightarrow \infty} \tilde{X}(k) = 0$ and the system (9.11a) is stable.

Denote $w(k) = \xi(k)\frac{1}{g(k)}$, $z(k) = \tilde{X}(k)\frac{1}{g(k)}$. According to the definition of $g(k)$, it follows from (9.11a)-(9.11b) that

$$\begin{cases} w(k+1) = \gamma^{-1}w(k) - h \sum_{i=0}^d L_i(k)\gamma^{-i-1}w(k-i) + h \sum_{i=0}^d L_i(k)\gamma^{-i-1}z(k-i), & (9.12a) \\ z(k+1) = \gamma^{-1} \{ \bar{e}(k) - q[\bar{e}(k)] \}, & (9.12b) \end{cases}$$

where

$$\bar{e}(k) = e(k)\frac{1}{g(k)} = z(k) + h \sum_{i=0}^d L_i(k)\gamma^{-i}z(k-i) - h \sum_{i=0}^d L_i(k)\gamma^{-i}w(k-i). \tag{9.13}$$

Note that $\gamma z(k+1)$ is equal to the quantization error of $\bar{e}(k)$. If $\bar{e}(k)$ never makes the quantizers saturate, then we have $\|z(k)\|_\infty \leq \frac{1}{2\gamma}$. If the delay system (9.12a)

is further stable, $\sup_{k \geq 0} \|w(k)\|_\infty$ must be bounded, which leads to (9.9) when $g(k)$ satisfies that $\lim_{k \rightarrow \infty} g(k) = 0$. Actually, if any of the quantizers is saturated, the problem becomes complicated since the boundedness of $z(k)$ as well as $w(k)$ cannot be guaranteed. Then, we note that the average consensus is solved if the following problem can be solved:

Problem 9.1. *Find the consensus gain function h , scaling function parameters g_0 , γ and quantization level $2K + 1$ such that*

1. *system (9.12a) is stable;*
2. *$\bar{e}(k)$ never makes the quantizers saturate.*

In the rest of this chapter, we shall consider the above problem instead.

Remark 9.2. *The two systems (9.12a) and (9.12b) play a key role in the average consensus problem. However, the two systems are coupled and (9.12b) is a non-linear system. Moreover, it is very challenging to analyze the stability of system (9.12a) since it is a time-varying system with multiple state delays. If we apply the augmentation approach to system (9.12a), we have the augmented system of the form:*

$$W(k+1) = \begin{bmatrix} \frac{I-hL_0(k)}{\gamma} & \frac{-hL_1(k)}{\gamma^2} & \cdots & \frac{-hL_d(k)}{\gamma^{d+1}} \\ I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \end{bmatrix} W(k) + \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} h \sum_{i=0}^d L_i(k) \gamma^{-i-1} z(k-i),$$

where $W(k) = \text{col}\{w(k), w(k-1), \dots, w(k-d)\}$. However, it is not clear how to choose h such that the augmented system is stable. In this chapter, we will introduce an auxiliary system. Based on the analysis of the auxiliary system, we can show how to select the control gain h such that system (9.12a) is stable.

Consider the following auxiliary system,

$$\begin{cases} w_a(k+1) = A(k)w_a(k) + H(k), & k \geq d, \\ H(k) = \sum_{i=1}^d B_i(k)H(k-i) + h \sum_{i=0}^d L_i(k)\gamma^{-i-1}z(k-i), \end{cases} \quad (9.14a)$$

$$(9.14b)$$

with initial conditions $w_a(d) = w(d)$, $H(0) = \dots = H(d-1) = 0$, where $A(k)$ and $B_i(k)$, $i = 1, 2, \dots, d$ satisfy

$$A(k) + B_1(k) = \gamma^{-1}(I - hL_0(k)), \quad (9.15)$$

$$B_1(k)A(k-1) - B_2(k) = \gamma^{-2}hL_1(k), \quad (9.16)$$

$$\vdots$$

$$B_{d-1}(k)A(k-d+1) - B_d(k) = \gamma^{-d}hL_{d-1}(k), \quad (9.17)$$

$$B_d(k)A(k-d) = \gamma^{-d-1}hL_d(k). \quad (9.18)$$

By simple calculation, it can be shown that $w(k) = w_a(k)$ for all $k \geq d$. Note that given the initial conditions $A(i)$, $i = 0, 1, \dots, d-1$, the matrices $A(k)$ and $B_i(k)$, $i = 1, 2, \dots, d$ can be calculated iteratively by (9.15)-(9.18). Define

$$\rho_A(t, \theta) = t^2 \frac{(1-\theta)^{-d} - 1 - d\theta}{\theta(\theta + 1 + d\theta - (1-\theta)^{-d})}, \quad (9.19)$$

$$\rho_B(t, \theta, l) = \sqrt{N} (t^2 + \theta\rho_A(t)) \frac{(1-\theta)^{-d} - 1 - d\theta}{\theta^2(1-\theta)^{ld}}, \quad (9.20)$$

where

$$\theta \in \{a : a + 1 + da > (1-a)^{-d}, a \in (0, 1)\}. \quad (9.21)$$

It is obvious that the set in (9.21) is not empty for any given positive integer d . For the auxiliary system (9.14b), we have the following lemma, whose proof can be found in the Appendix.

Lemma 9.1. *Consider system (9.14a)-(9.14b). Denote*

$$\Delta_A(k) = \gamma A(k) - I + hL(k), \quad \Delta_{B_j}(k) = B_j(k) - \gamma^{-j}h \sum_{i=j}^d L_i(k), \quad j = 1, \dots, d, \quad (9.22)$$

where $L(k) = \sum_{i=0}^d L_i(k)$. For any given constant θ satisfying (9.21), constant $\bar{\lambda}$ satisfying $\bar{\lambda} \geq \sup_k \lambda(L(k))$, if

$$\|\Delta_A(k)\|_2 < \rho_A(\bar{\lambda}, \theta)h^2, \quad k = 0, 1, \dots, d-1, \quad (9.23)$$

and $0 < h < h_1(\bar{\lambda}, \theta)$, where

$$h_1(\bar{\lambda}, \theta) = \frac{2\theta}{\sqrt{\bar{\lambda}^2 + 4\rho_A(\bar{\lambda}, \theta)\theta} + \bar{\lambda}}, \quad (9.24)$$

then $\forall k \geq d$, $A(k)$ is invertible and

$$\|\Delta_A(k)\|_2 < \rho_A(\bar{\lambda}, \theta)h^2. \quad (9.25)$$

Furthermore, given any $l > 0$, if $(1 - \theta)^l < \gamma < 1$, then

$$\|\Delta_{B_j}(k)\|_\infty < \rho_B(\bar{\lambda}, \theta, l)h^2, \quad j = 1, \dots, d. \quad (9.26)$$

In what follows, we shall consider the problem in two cases separately, including constant transmission delays and time-varying transmission delays.

Case 1: Constant delays

In this case, we shall consider that all the communication delays are time-invariant.

It is obvious that

$$\sum_{l=0}^d L_l(k) = L, \quad \forall k \geq d.$$

Denote λ_i the i -th smallest eigenvalue of L . Before giving the main result, the following lemma is needed.

Lemma 9.2. *Consider system (9.12a)-(9.12b). Define*

$$M_1(h, \gamma, g_0) = \frac{1}{2\gamma} + \left(\frac{1}{2\gamma} + f_1(h, \gamma, d) \right) 2hD^* \frac{\gamma^{-d} - 1}{\gamma^{-1} - 1}, \quad (9.27)$$

where f_1 is given by

$$f_1(h, \gamma, k) = p_1^k \frac{2\sqrt{N}C_x}{g_0} + \frac{h\sqrt{N}\lambda_N(\gamma^{-d} - 1)(p_1^k - 1)}{2\gamma(1 - \gamma)(p_1 - 1)}, \quad p_1 = \gamma^{-1} + h\lambda_N \frac{\gamma^{-d} - 1}{1 - \gamma}. \quad (9.28)$$

If $g_0 \geq 2\gamma C_x$, and $K \geq \lfloor M_1(h, \gamma, g_0) + \frac{1}{2} \rfloor$, then

$$\|w(k)\|_2 \leq f_1(h, \gamma, k), \quad \|z(k)\|_\infty \leq \frac{1}{2\gamma}, \quad \forall k = 0, 1, \dots, d. \quad (9.29)$$

Proof. Set $k = 0$, then we have

$$\|w(0)\|_2 \leq \sqrt{N} \|w(0)\|_\infty = \frac{\sqrt{N} \|\xi(0)\|_\infty}{g_0} \leq \frac{2\sqrt{N}C_x}{g_0} = f_1(h, \gamma, 0), \quad (9.30)$$

which is the first inequality of (9.29). The last inequality of (9.30) is due to the fact that $\|\xi(0)\|_\infty \leq 2C_x$. According to the definition of $z(k)$, we have

$$\|z(0)\|_\infty = \left\| \frac{\tilde{X}(0)}{g_0} \right\|_\infty = \left\| \frac{X(0)}{g_0} \right\|_\infty \leq \frac{C_x}{g_0} \leq \frac{1}{2\gamma},$$

which is the second inequality of (9.29).

Next, we assume that (9.29) is satisfied when $k = 0, 1, \dots, r$ with $r < d$. It is clear that $f_1(h, \gamma, k)$ is an increasing function due to that $p_1 > 1$. According to the definition of $\bar{e}(k)$ we have that

$$\begin{aligned} \|\bar{e}(r)\|_\infty &\leq \left\| z(r) + h \sum_{i=0}^r L_i(r) \gamma^{-i} z(r-i) - h \sum_{i=0}^r L_i(r) \gamma^{-i} w(r-i) \right\|_\infty \\ &\leq \left(1 + h \sum_{i=0}^r \|L_i(r)\|_\infty \gamma^{-i} \right) \frac{1}{2\gamma} + h \sum_{i=0}^r \|L_i(r)\|_\infty \gamma^{-i} f_1(h, \gamma, r) \\ &\leq \left(1 + h \|L\|_\infty \frac{\gamma^{-r-1} - 1}{\gamma^{-1} - 1} \right) \frac{1}{2\gamma} + h \|L\|_\infty f_1(h, \gamma, r) \frac{\gamma^{-r-1} - 1}{\gamma^{-1} - 1} \\ &\leq M_1(h, \gamma, g_0). \end{aligned}$$

Because $K \geq \lfloor M_1(h, \gamma, g_0) + \frac{1}{2} \rfloor$, we know that $\bar{e}(r)$ does not make the quantizer saturate, which, according to (9.12b), leads to that $\|z(r+1)\|_\infty \leq \frac{1}{2\gamma}$.

From (9.12a), we have

$$\begin{aligned}
\|w(r+1)\|_2 &\leq \gamma^{-1} \left\| w(r) - h \sum_{i=0}^r L_i(r) \gamma^{-i} w(r-i) \right\|_2 + \gamma^{-1} \left\| h \sum_{i=0}^r L_i(r) \gamma^{-i} z(r-i) \right\|_2 \\
&\leq \gamma^{-1} \left(1 + h \sum_{i=0}^r \|L_i(r)\|_2 \gamma^{-i} \right) f_1(h, \gamma, r) + \gamma^{-1} \frac{\sqrt{N}}{2\gamma} h \sum_{i=0}^r \|L_i(r)\|_2 \gamma^{-i} \\
&\leq p_1 f_1(h, \gamma, r) + \frac{\sqrt{N}}{2\gamma} h \|L\|_2 \frac{\gamma^{-r-1} - 1}{1 - \gamma} \\
&= f_1(h, \gamma, r+1),
\end{aligned}$$

which means that (9.29) holds for $k = r + 1$. By induction, the proof is completed.

□

Define

$$\bar{h}_\theta = h_1(\lambda_N, \theta) \wedge \frac{\lambda_2}{\rho_A(\lambda_N, \theta)} \wedge \left(1 + 2D^*(1 - \theta)^{-d} + \frac{\rho_B(\lambda_N, \theta, 1)}{4} \right)^{-d}, \quad (9.31)$$

where h_1 is defined in (9.24), and

$$\bar{\gamma}_h = 1 - \lambda_2 h + \rho_A(\lambda_N, \theta) h^2, \quad (9.32)$$

$$K_1(h, \gamma, g_0) = \frac{1}{2\gamma} + \left(\frac{1}{2\gamma} + \bar{w}(h, \gamma) \right) 2hD^* \frac{\gamma^{-d-1} - 1}{\gamma^{-1} - 1}, \quad (9.33)$$

with $\bar{w}$ defined as

$$\bar{w}(h, \gamma) = f_1(h, \gamma, d) \vee \frac{h^{\frac{1}{d}} D^* \sqrt{N} (\gamma^{-d-1} - 1)}{(1-\gamma)(1-\beta)(\gamma - \bar{\gamma}_h)}, \quad \beta = h^{\frac{1}{d}} \vee h^{\frac{1}{d}} \frac{2D^*(1-\theta)^{-d} + \rho_B(\lambda_N, \theta, 1)/4}{1 - h^{\frac{1}{d}}}. \quad (9.34)$$

Now we are in the position to give the main result.

Theorem 9.1. *Consider system (9.1) with the Assumptions 9.1-9.4 satisfied. Apply the protocol (9.5) with the encoder (9.2) and decoder (9.3). For any given constant θ satisfying (9.21), if the control parameters h, γ, g_0 satisfy*

$$h \in (0, \bar{h}_\theta), \quad \gamma \in (\bar{\gamma}_h, 1), \quad g_0 \geq 2\gamma C_x, \quad (9.35)$$

and the number of quantization level satisfies

$$K \geq \lfloor K_1(h, \gamma, g_0) + \frac{1}{2} \rfloor, \quad (9.36)$$

we have the following results:

1) The average consensus problem is solved, and

$$\lim_{h \rightarrow 0^+, \gamma \rightarrow 1^-} K_1(h, \gamma, g_0) = \frac{1}{2}, \quad (9.37)$$

i.e. the lower bound of the number of transmitting bits is 1.

2) The consensus error $\xi(k)$ satisfies

$$\limsup_{k \rightarrow \infty} \frac{\|\xi(k)\|_2}{\gamma^k} \leq \bar{w}g_0, \quad (9.38)$$

and $r_{\text{asym}} \leq \gamma$, where

$$r_{\text{asym}} = \sup_{\xi(0) \neq 0} \lim_{k \rightarrow \infty} \|X(k) - J_N X(0)\|_2^{1/k} \quad (9.39)$$

is the convergence rate of the average consensus defined in [131].

Proof. We shall consider the stability of system (9.14a)-(9.14b) with initial condition of $A(k)$ satisfying

$$\|\Delta_A(k)\|_2 < \rho_A(\lambda_N, \theta)h^2, \quad k = 0, 1, \dots, d-1,$$

where λ_N is the largest eigenvalue of the Laplacian matrix, θ satisfies (9.21). Then, according to (9.31), (9.35) and Lemma 9.1, we have $\|\Delta_A(k)\|_2 < \rho_A(\lambda_N, \theta)h^2$, $\forall k \geq 0$. Note that $(I - J_N)\xi(k) = (I - J_N)^2 X(k) = (I - J_N)X(k) = \xi(k)$. According to the definition of $w(k)$, it is clear that $(I - J_N)w(k) = w(k)$. Since $w_a(k) = w(k)$ for all $k \geq d$, then system (9.14a) can be written as

$$w_a(k+1) = \bar{A}(k)w_a(k) + H(k), \quad (9.40)$$

where $\bar{A}(k) = A(I - J_N) \triangleq \gamma^{-1}(I - hL - J_N + \Delta_{\bar{A}}(k))$ with $\Delta_{\bar{A}}(k) = \Delta_A(k)(I - J_N)$. Then we have

$$\|\Delta_{\bar{A}}(k)\|_2 \leq \|\Delta_A(k)\|_2 \|I - J_N\|_2 = \|\Delta_A(k)\|_2 < \rho_A(\lambda_N, \theta)h^2.$$

Note that the eigenvalues of $I - hL - J_N$ are 0, $1 - \lambda_2, \dots, 1 - \lambda_N$. Since the communication graph is connected, from Lemma A.2, we know that $\lambda_2 > 0$. According to (9.32) and (9.35), we have

$$\|\bar{A}(k)\|_2 \leq \gamma^{-1}(\|I - hL - J_N\|_2 + \|\Delta_A(k)\|_2) \leq \frac{\bar{\gamma}_h}{\gamma} < 1. \quad (9.41)$$

(9.41) implies the stability of system (9.40) and equivalently that of (9.14a).

Next, we shall analyze the stability of (9.14b). We rewrite system (9.14b) in the following augmented form

$$\bar{H}(k) = B(k)\bar{H}(k-1) + \Gamma h \sum_{i=0}^d L_i(k) \gamma^{-i-1} z(k-i), \quad (9.42)$$

where

$$B(k) = \begin{bmatrix} B_1(k) & B_2(k) & B_3(k) & \cdots & B_d(k) \\ I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \end{bmatrix},$$

$$\Gamma = [I \ 0 \ \cdots \ 0]', \quad \bar{H}(k) = [H'(k) \ \cdots \ H'(k-d+1)]'.$$

By introducing the linear transformation

$$T = \begin{cases} \text{diag}\{I, h^{1/d}I, h^{2/d}I, \dots, h^{d-1/d}I\}, & d \neq 1, \\ I, & d = 1, \end{cases}$$

we have

$$\begin{aligned}\bar{B}(k) &\triangleq TB(k)T^{-1} \\ &= \begin{bmatrix} B_1(k) & \frac{B_2(k)}{h^{1/d}} & \frac{B_3(k)}{h^{2/d}} & \cdots & \frac{B_d(k)}{h^{(d-1)/d}} \\ h^{1/d}I & & & & 0 \\ & h^{1/d}I & & & 0 \\ & & \ddots & & \vdots \\ & & & h^{1/d}I & 0 \end{bmatrix}. \end{aligned} \quad (9.43)$$

By (9.31), it follows that $\bar{\gamma}_h > 1 - \lambda_2 h - \rho_A(\lambda_N, \theta)h^2 > 1 - \theta$. According to Lemma 9.1 and (9.35), we know that $\|\Delta_{B_j}(k)\|_\infty < \rho_B(\lambda_N, \theta, 1)h^2$. Moreover, by considering Lemma A.2 and Lemma 9.1, it follows that

$$\|B_j(k)\|_\infty = \left\| \gamma^{-j}h \sum_{i=j}^d L_i(k) + \Delta_{B_j}(k) \right\|_\infty \leq 2h\gamma^{-j}D^* + \rho_B(\lambda_N, \theta, 1)h^2,$$

where D^* is the degree of the graph. Then we calculate the norm of $\bar{B}$ based on (9.43) as follows

$$\begin{aligned}\|\bar{B}(k)\|_\infty &\leq h^{\frac{1}{d}} \vee \sum_{i=1}^d h^{\frac{1-i}{d}} \|B_i(k)\|_\infty \\ &\leq h^{\frac{1}{d}} \vee \sum_{i=1}^d h^{\frac{1-i}{d}} (2h\gamma^{-i}D^* + \rho_B(\lambda_N, \theta, 1)h^2) \\ &\leq h^{\frac{1}{d}} \vee \left[2h^{\frac{d+1}{d}}D^* \sum_{i=1}^d \left(h^{\frac{1}{d}}\gamma\right)^{-i} + \rho_B(\lambda_N, \theta, 1)h^2 \sum_{i=1}^d h^{\frac{1-i}{d}} \right] \\ &< h^{\frac{1}{d}} \vee \left[\frac{2h^{\frac{1}{d}}D^*\gamma^{-d}}{1 - \gamma h^{\frac{1}{d}}} + \rho_B(\lambda_N, \theta, 1)h^{\frac{d+1}{d}} \frac{1-h}{1-h^{\frac{1}{d}}} \right] \\ &< h^{\frac{1}{d}} \vee \left[h^{\frac{1}{d}} \frac{2D^*(1-\theta)^{-d} + \rho_B(\lambda_N, \theta, 1)/4}{1-h^{\frac{1}{d}}} \right] = \beta. \end{aligned} \quad (9.44)$$

The last inequality is due to that $1 - \gamma h^{\frac{1}{d}} > 1 - h^{\frac{1}{d}}$ and $h(1-h) \leq \frac{1}{4}$. From the above inequality we can see that $\beta < 1$ if and only if $h < \left(1 + 2D^*(1-\theta)^{-d} + \frac{\rho_B}{4}\right)^{-d}$ which is satisfied according to (9.31) and (9.35). On the other hand, $\beta < 1$ implies

the stability of system (9.42). In a word, (9.35) implies the stability of system (9.14a)-(9.14b).

Last, we shall prove that when $K_1(h, \gamma, g_0)$ is large enough the quantizer will never be saturated for any K satisfying (9.36). This can be done by induction. It is easy to check that $K_1(h, \gamma, g_0) > M_1(h, \gamma, g_0)$, where M_1 is as defined in (9.27), which means that $\|\bar{e}(k)\|_\infty \leq K_1(h, \gamma, g_0)$ for $k = 0, 1, \dots, d$. Suppose that

$$\|\bar{e}(k)\|_\infty \leq K_1(h, \gamma, g_0), \quad \forall k = 0, 1, \dots, s,$$

where $s \geq d$. According to (9.12b) we see that $\|z(k)\|_\infty \leq \frac{1}{2\gamma}$ for all $k = 0, 1, \dots, s+1$, $s \geq d$. Considering (9.42) with initial condition $\bar{H}(0) = 0$, we have

$$\|\Pi_{i,j}^B\|_\infty \leq \|T\|_\infty \|T^{-1}\|_\infty \|\Pi_{i,j}^{\bar{B}}\|_\infty \leq h^{\frac{1-d}{d}} \beta^{i-j+1},$$

and

$$\begin{aligned} \|\bar{H}(s)\|_\infty &= \left\| \sum_{j=1}^s \Pi_{s,j+1}^B \Gamma h \sum_{i=0}^d L_i(j) \gamma^{-i-1} z(j-i) \right\|_\infty \\ &\leq h^{\frac{1}{d}} \sum_{j=1}^s \beta^{s-j} \|\Gamma\|_\infty \sum_{i=0}^d \gamma^{-i-1} \|L_i(j)\|_\infty \|z(j-i)\|_\infty \\ &\leq h^{\frac{1}{d}} \sum_{j=0}^{s-1} \beta^j \left(\sum_{i=0}^d \gamma^{-i-1} \|L\|_\infty \frac{1}{2\gamma} \right) \\ &< h^{\frac{1}{d}} D^* \frac{\gamma^{-d-1} - 1}{\gamma(1-\gamma)(1-\beta)}. \end{aligned} \tag{9.45}$$

Based on (9.14a), (9.45) and the fact $\|v\|_\infty \leq \|v\|_2 \leq \sqrt{n}\|v\|_\infty$ for any $n \times 1$ vector v , we have

$$\begin{aligned} \|w_a(s+1)\|_2 &= \left\| \Pi_{s,d}^{\bar{A}} w_a(d) + \sum_{i=d}^s \Pi_{s,i+1}^{\bar{A}} H(i) \right\|_2 \\ &\leq \left(\frac{\bar{\gamma}_h}{\gamma} \right)^{s-d+1} \|w_a(d)\|_2 + \sum_{i=d}^s \left(\frac{\bar{\gamma}_h}{\gamma} \right)^{s-i} \sqrt{N} \|H(i)\|_\infty \end{aligned}$$

$$\begin{aligned}
&\leq \left(\frac{\bar{\gamma}_h}{\gamma}\right)^{s-d+1} \|w_a(d)\|_2 + \left[1 - \left(\frac{\bar{\gamma}_h}{\gamma}\right)^{s-d+1}\right] \frac{h^{\frac{1}{d}} D^* \sqrt{N} (\gamma^{-d-1} - 1)}{(1-\gamma)(1-\beta)(\gamma - \bar{\gamma}_h)} \\
&\leq \bar{w},
\end{aligned} \tag{9.46}$$

where $\bar{w}$ is defined in (9.34). The last inequality in (9.46) is obtained from Lemma 9.2 and $w_a(d) = w(d)$. Since $w_a(k) = w(k)$ for all $k \geq d$, we have $\|w(k)\|_2 \leq \bar{w}$ for all $k = 0, 1, \dots, s+1$. Then according to the definition of $\bar{e}(k)$ in (9.13), it follows from (9.46) that

$$\begin{aligned}
\|\bar{e}(s+1)\|_\infty &= \left\| z(s+1) + h \sum_{i=0}^d L_i(s+1) \gamma^{-i} z(s+1-i) \right. \\
&\quad \left. - h \sum_{i=0}^d L_i(s+1) \gamma^{-i} w(s+1-i) \right\|_\infty \\
&\leq \frac{1 + h \sum_{i=0}^d \|L_i(s+1)\|_\infty \gamma^{-i}}{2\gamma} + h\bar{w} \sum_{i=0}^d \|L_i(s+1)\|_\infty \gamma^{-i} \\
&\leq \frac{1}{2\gamma} + \left(\frac{1}{2\gamma} + \bar{w}\right) 2hD^* \frac{\gamma^{-d-1} - 1}{\gamma^{-1} - 1} \\
&= K_1(h, \gamma, g_0).
\end{aligned} \tag{9.47}$$

By induction, we know that $\|\bar{e}(k)\|_\infty \leq K_1(h, \gamma, g_0)$ and $\|w(k)\|_2 \leq \bar{w}$ for all $k \geq 0$, that is to say, the quantizer will never be saturated if (9.36) is satisfied. By the definition of $w(k)$ and $0 < \gamma < 1$, (9.9) is obtained.

From (9.28), (9.32) and (9.44), we have

$$\begin{aligned}
\lim_{h \rightarrow 0^+, \gamma \rightarrow 1^-} |f(h, \gamma, d)| &< +\infty, \quad \lim_{\gamma \rightarrow 1^-} \frac{\gamma^{-d-1} - 1}{\gamma^{-1} - 1} = d+1, \\
\lim_{h \rightarrow 0^+, \gamma \rightarrow 1^-} \frac{h}{\gamma - \bar{\gamma}_h} &= \frac{1}{\lambda_2}, \quad \lim_{h \rightarrow 0^+} \beta = 0.
\end{aligned}$$

It follows that

$$\lim_{h \rightarrow 0^+, \gamma \rightarrow 1^-} K_1(h, \gamma, g_0) = \frac{1}{2},$$

which is (9.37). So far, we have proved 1).

It is noted that

$$\|\xi(k)\|_2 = \|g(k)w(k)\|_2 \leq \bar{w}g_0\gamma^k,$$

which leads to (9.38). According to (9.8) we arrive at

$$\begin{aligned} \lim_{k \rightarrow \infty} \|X(k) - J_N X(0)\|_2^{1/k} &= \lim_{k \rightarrow \infty} \exp \left\{ \frac{1}{k} \ln \|\xi(k)\|_2 \right\} \\ &\leq \exp \left\{ \lim_{k \rightarrow \infty} \frac{\ln \|\xi(k)\|_2}{k} \right\} \\ &\leq \exp \left\{ \lim_{k \rightarrow \infty} \frac{\ln (\bar{w}g_0\gamma^k)}{k} \right\} \\ &= \gamma. \end{aligned}$$

By this and the definition of r_{asym} we get (9.39). $\square$

Remark 9.3. We have provided the ranges of the consensus gain h and scaling parameter γ such that $w(k)$ and $z(k)$ are uniformly bounded, and consequently, the consensus error $\xi(k)$ converges to 0 at the exponential rate $O(\gamma^k)$. We have worked out a lower bound of the quantization level K , which is a continuous function of h , γ and g_0 . From (9.36) we can see that the bit rate can reach 1 when $K_1 < 1$. On the other hand, K_1 can approach $\frac{1}{2}$ when $h \rightarrow 0^+$ and $\gamma \rightarrow 1^-$. Hence there is a trade-off between the consensus rate and communication data rate. A lower communication data rate requires a larger γ which leads to a lower consensus convergence rate $O(\gamma^k)$. On the other hand, a smaller γ can guarantee a faster convergence but, according to (9.33), (9.34) and (9.36), a higher data rate is required for each channel. This result is coincident with the delay-free case [64].

Remark 9.4. When we design the parameters h , γ , g_0 and calculate the lower bound of quantization level, some knowledge is required, such as the positive lower bound of λ_2 , the upper bound of D^{*1} . In fact, this can be estimated for some special graphs. For example, for expander networks, if the communication graph is a d -regular c -expander, we have $\lambda_2 \geq d - \sqrt{d^2 - c^2}$, $D^* = d$.

The above result for constant time delay case provides a new way to solve such

¹The upper bound of λ_N can be estimated by $\lambda_N \leq 2D^*$.

complicated consensus problem. Based on this result, in the next subsection, we shall consider a more general case, in which the communication delays are time-varying.

Case 2: Time-varying delays

We shall consider that the delays in the transmission channels are time-varying. Note that the summation $\sum_{l=0}^d L_l(k) \triangleq L(k)$ is no longer a constant matrix due to the time-varying delays. It is much more complicated to analyze the stability of the system than the one with constant delays. Nevertheless, $L(k)$ has the following properties

$$L \leq \sum_{l=0}^d L(k+l), \quad L(k) \leq dL. \quad (9.48)$$

In fact, due to that the time delays are bounded by d , a packet in each channel will be received within d steps, which implies the first inequality in (9.48). On the other hand, the receiver can receive at most d packets from each channel at one time, which yields the second inequality in (9.48). Similar to Lemma 9.2, we have the following lemma. The proof is similar to that of Lemma 9.2 by considering (9.48) and thus is omitted.

Lemma 9.3. *Consider system (9.12a)-(9.12b). Define*

$$M_2(h, \gamma, g_0) = \frac{1}{2\gamma} + \left(\frac{1}{2\gamma} + f_2(h, \gamma, d) \right) 2hdD^* \frac{\gamma^{-d} - 1}{\gamma^{-1} - 1}, \quad (9.49)$$

where f_2 is given by

$$f_2(h, \gamma, k) = p_2^k \frac{2\sqrt{N}C_x}{g_0} + \frac{hd\sqrt{N}\lambda_N (\gamma^{-d} - 1) (p_2^k - 1)}{2\gamma (1 - \gamma) (p_2 - 1)}, \quad p_2 = \gamma^{-1} + h\lambda_N d \frac{\gamma^{-d} - 1}{1 - \gamma}.$$

If $g_0 \geq 2\gamma C_x$ and $K \geq \lfloor M_2(h, \gamma, g_0) + \frac{1}{2} \rfloor$, then the following inequalities hold $\forall k = 0, 1, \dots, d$

$$\|w(k)\|_2 \leq f_2(h, \gamma, k), \quad \|z(k)\|_\infty \leq \frac{1}{2\gamma}. \quad (9.50)$$

Define

$$\bar{h}_{\theta,1} = h_1(\lambda_N d, \theta) \wedge \frac{\lambda_2}{\bar{\rho}_A(\theta)} \wedge \left(1 + 2dD^*(1-\theta)^{-1} + \frac{\rho_B(\lambda_N d, \theta, (d+1)^{-1})}{4} \right)^{-d}, \quad (9.51)$$

with

$$\bar{\rho}_A(\theta) \triangleq (d+1)\rho_A(\lambda_N d, \theta) + \frac{(1+\theta)^{d+1}-1-(d+1)\theta}{\theta^2} \left(\frac{\sqrt{d^2\lambda_N^2+4\alpha\rho_A(\lambda_N d, \theta)+\lambda_N d}}{2} \right)^2, \quad (9.52)$$

and

$$\bar{\gamma}_{h,1} = 1 - \lambda_2 h + \bar{\rho}_A(\theta) h^2, \quad (9.53)$$

$$K_2(h, \gamma, g_0) = \frac{1}{2\gamma} + \left(\frac{1}{2\gamma} + \bar{w}_1(h, \gamma) \right) 2hdD^* \frac{\gamma^{-d-1} - 1}{\gamma^{-1} - 1}, \quad (9.54)$$

with

$$\begin{aligned} \bar{w}_1(h, \gamma) \triangleq & \left[\left(\frac{1 + \rho_A(\lambda_N d, \theta) h^2}{\gamma} \right)^d f_2(h, \gamma, d) \right] \vee \left[\frac{\left(\frac{1 + \rho_A(\lambda_N d, \theta) h^2}{\gamma} \right)^{d+1} - 1}{\left(\frac{1 + \rho_A(\lambda_N d, \theta) h^2}{\gamma} \right) - 1} \right. \\ & \left. \cdot \frac{h^{\frac{1}{d}} d D^* \sqrt{N} (\gamma^{-d-1} - 1)}{(1-\gamma)(1-\beta_1)(\gamma - \gamma^{-d} \bar{\gamma}_{h,1})} \right], \end{aligned} \quad (9.55)$$

$$\beta_1 = h^{\frac{1}{d}} \vee \frac{2dD^*(1-\theta)^{-1} + \rho_B(\lambda_N d, \theta, (d+1)^{-1})/4}{h^{-\frac{1}{d}} - 1}. \quad (9.56)$$

Then we have the following result.

Theorem 9.2. *Consider system (9.1) with time-varying transmission delays. Under Assumptions 9.1-9.4, we apply the protocol (9.5) with encoder (9.2) and decoder (9.3). For any given $\theta > 0$ satisfying (9.21), if h, γ, g_0 satisfy*

$$h \in (0, \bar{h}_{\theta,1}), \quad \gamma \in (\sqrt[d+1]{\bar{\gamma}_{h,1}}, 1), \quad g_0 \geq 2\gamma C_x, \quad (9.57)$$

and the number of quantization level satisfies

$$K \geq \lfloor K_2(h, \gamma, g_0) + \frac{1}{2} \rfloor, \quad (9.58)$$

we have the following result:

1) The average consensus problem is solved. In particular, $K_2(h, \gamma, g_0)$ satisfies

$$\lim_{h \rightarrow 0^+, \gamma \rightarrow 1^-} K_2(h, \gamma, g_0) = \frac{1}{2}, \quad (9.59)$$

i.e. the lower bound of the number of transmitting bits is 1.

2) We have $r_{\text{asym}} \leq \gamma$, and the consensus error $\xi(k)$ satisfies

$$\limsup_{k \rightarrow \infty} \frac{\|\xi(k)\|_2}{\gamma^k} \leq \bar{w}g_0. \quad (9.60)$$

Proof. The proof is partly similar to that of Theorem 9.1, so only the different part will be pointed out. We consider system (9.14a)-(9.14b) with initial condition of $A(k)$ satisfying

$$\|\Delta_A(k)\|_2 < \rho_A(\lambda_N d, \theta)h^2, \quad \forall k = 0, 1, \dots, d-1,$$

where θ satisfies (9.21). By (9.48), we can see that $\sup_{k \geq 0} \lambda(L(k)) \leq \lambda_N d$. Then, according to Lemma 9.1 and (9.57), we have that $\|\Delta_A(k)\|_2 < \rho_A(\lambda_N d, \theta)h^2$ for all $k \geq 0$. System (9.14a) can be rewritten as

$$w_a(k+1) = \bar{A}(k)w_a(k) + H(k), \quad (9.61)$$

where

$$\bar{A}(k) = A(k)(I - J_N) = \gamma^{-1}(I - hL(k) - J_N + \Delta_{\bar{A}}(k)), \quad (9.62)$$

with $\Delta_{\bar{A}}(k) \triangleq \Delta_A(k)(I - J_N)$. It is clear that $\|\Delta_{\bar{A}}(k)\|_2 < \rho_A(\lambda_N d, \theta)h^2$. According to (9.57), by simple calculation, it is shown that

$$\lambda_N d + \rho_A(\lambda_N d, \theta)h < \frac{\sqrt{d^2 \lambda_N^2 + 4\bar{\rho}_A \theta} + \lambda_N d}{2}. \quad (9.63)$$

By defining $g(k) = \max_{j=0,\dots,d} \|hL(k+j) - \Delta_{\bar{A}}(k+j)\|_2$, it is easy to check that

$$g(k) \leq \lambda_N dh + \rho_A(\lambda_N d, \theta) h^2 < \theta, \quad (9.64)$$

and therefore

$$\frac{(1+g(k))^{d+1} - 1 - (d+1)g(k)}{g^2(k)} < \frac{(1+\theta)^{d+1} - 1 - (d+1)\theta}{\theta^2}. \quad (9.65)$$

Then we have

$$\begin{aligned} \|\Pi_{k+d,k}^{\gamma \bar{A}}\|_2 &= \gamma^{d+1} \|\Pi_{k+d,k}^{\bar{A}}\|_2 \\ &\leq \|(I - J_N - hL(k+d) + \Delta_{\bar{A}}(k+d)) \cdots (I - J_N - hL(k) + \Delta_{\bar{A}}(k))\|_2 \\ &\leq \left\| I - J_N - h \sum_{i=0}^d L(k+i) \right\|_2 + (d+1)\rho_A(\lambda_N d, \theta) h^2 \\ &\quad + [(1+g(k))^{d+1} - 1 - (d+1)g(k)] \\ &< \left\| I - J_N - h \sum_{i=0}^d L(k+i) \right\|_2 + \left[(d+1)\rho_A(\lambda_N d, \theta) \right. \\ &\quad \left. + \frac{(1+\theta)^{d+1} - 1 - (d+1)\theta}{\theta^2} \left(\frac{\sqrt{d^2 \lambda_N^2 + 4\rho_A(\lambda_N d, \theta)\theta} + \lambda_N d}{2} \right)^2 \right] h^2 \\ &\leq \bar{\gamma}_{h,1}, \end{aligned} \quad (9.66)$$

where $\bar{\gamma}_{h,1}$ is as defined in (9.53). The second last inequality is due to (9.63) and (9.65). From (9.66) we can see that $^{d+1}\sqrt{\bar{\gamma}_{h,1}} < \gamma < 1$ implies $\|\Pi_{k+d,k}^{\bar{A}}\|_2 < 1$ which guarantees the stability of system (9.61). By considering (9.64), we know that $\theta > \lambda_N dh + \rho_A(\lambda_N d, \theta) h^2$, which leads to that $\bar{\gamma}_{h,1} > 1 - \theta$. According to Lemma 9.1, we have

$$\|\Delta_{B_j}(k)\|_\infty < \rho_B(\lambda_N d, \theta, (d+1)^{-1}) h^2,$$

where $\Delta_{B_j}(k)$ is as defined in (9.22). Again we introduce the augmented system (9.42) of (9.14b). Then by applying the same manipulation as in (9.44), for matrix $\bar{B}(k)$ defined in (9.43), we have $\|\bar{B}(k)\|_\infty < \beta_1$, where β_1 is as defined in (9.56).

Note that $\beta_1 < 1$ if and only if

$$h < \left(1 + 2dD^*(1 - \theta)^{-1} + \frac{\rho_B(\lambda_N d, \theta, (d+1)^{-1})}{4} \right)^{-d}. \quad (9.67)$$

(9.57) guarantees (9.67) and therefore the stability of system (9.42). In a word, h and γ satisfying (9.57) make sure that the system (9.14a)-(9.14b) is stable. Since $w_a(k) = w(k)$ for all $k \geq d$, we can conclude that system (9.12a) is stable which means we have solved the part 1) of problem 2.

Next, we need to determine K_2 such that the quantizer will not be saturated for any K satisfying (9.58). It is easy to check that $K_2(h, \gamma, g_0) > M_2(h, \gamma, g_0)$, where M_2 is as defined in (9.49), which means that $\|\bar{e}(k)\|_\infty \leq K_2(h, \gamma, g_0)$ for $k = 0, 1, \dots, d$. Suppose that

$$\|\bar{e}(k)\|_\infty \leq K_2(h, \gamma, g_0), \quad \forall k = 0, 1, \dots, s, \quad (9.68)$$

where $s \geq d$. According to (9.12b) we see that $\|z(k)\|_\infty \leq \frac{1}{2\gamma}$ for all $k = 0, 1, \dots, s+1$, $s \geq d$. Considering (9.42) with initial condition $\bar{H}(0) = 0$, along the same line of (9.45), we have

$$\|\bar{H}(s)\|_\infty < h^{\frac{1}{d}} d D^* \frac{\gamma^{-d-1} - 1}{\gamma(1 - \gamma)(1 - \beta_1)}. \quad (9.69)$$

From (9.62) we get

$$\|\bar{A}(k)\|_2 \leq \frac{\|I - hL(k) + \Delta_{\bar{A}}(k)\|_2 + \rho_A(\lambda_N d, \theta)h^2}{\gamma} \leq \frac{1 + \rho_A(\lambda_N d, \theta)h^2}{\gamma}, \quad (9.70)$$

which together with (9.66) results in that, for any $k \geq 0$, $m \geq 0$,

$$\begin{aligned} \left\| \Pi_{k+m, k+1}^{\bar{A}} \right\|_2 &\leq \left(\frac{1 + \rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^{m - \phi_{m,d}} \left(\frac{\bar{\gamma}_{h,1}}{\gamma^{d+1}} \right)^{\phi_{m,d}} \\ &\leq \left(\frac{1 + \rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^d \left(\frac{\bar{\gamma}_{h,1}}{\gamma^{d+1}} \right)^{\phi_{m,d}} \\ \sum_{i=0}^m \left\| \Pi_{k+i, k+1}^{\bar{A}} \right\|_2 &\leq \frac{1 - (\gamma^{-d-1} \bar{\gamma}_{h,1})^{\phi_{m,d}}}{1 - \gamma^{-d-1} \bar{\gamma}_{h,1}} \sum_{j=0}^d \left(\frac{1 + \rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^j \end{aligned}$$

$$\leq \frac{1 - (\gamma^{-d-1}\bar{\gamma}_{h,1})^{\phi_{m,d}} \left(\frac{1+\rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^{d+1} - 1}{1 - \gamma^{-d-1}\bar{\gamma}_{h,1}} \frac{\left(\frac{1+\rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^{d+1} - 1}{\left(\frac{1+\rho_A(\lambda_N d, \theta)h^2}{\gamma} \right) - 1}, \quad (9.71)$$

where $\phi_{m,d} \triangleq \lfloor \frac{m}{d+1} \rfloor$. Then we have

$$\begin{aligned} \|w_a(s+1)\|_2 &= \left\| \Pi_{s,d}^{\bar{A}} w_a(d) + \sum_{i=d}^s \Pi_{s,i+1}^{\bar{A}} H(i) \right\|_2 \\ &\leq \left(\frac{\bar{\gamma}_{h,1}}{\gamma^{d+1}} \right)^{\phi_{s-d+1,d}} \left(\frac{1 + \rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^d \|w_a(d)\|_2 + \frac{1 - \left(\frac{\bar{\gamma}_{h,1}}{\gamma^{d+1}} \right)^{\phi_{s-d+1,d}}}{1 - \gamma^{-d-1}\bar{\gamma}_{h,1}} \\ &\quad \times \frac{\left(\frac{1+\rho_A(\lambda_N d, \theta)h^2}{\gamma} \right)^{d+1} - 1}{\left(\frac{1+\rho_A(\lambda_N d, \theta)h^2}{\gamma} \right) - 1} \frac{h^{\frac{1}{d}} d D^* \sqrt{N} (\gamma^{-d-1} - 1)}{\gamma(1-\gamma)(1-\beta_1)} \\ &\leq \bar{w}, \end{aligned} \quad (9.72)$$

where $\bar{w}$ is defined in (9.55). The first inequality in (9.72) follows from (9.69) and (9.71), the second inequality is due to Lemma 9.3 and the definition of $\bar{w}$. Along the same line of (9.47), there holds that

$$\|\bar{e}(s+1)\|_\infty \leq K_2(h, \gamma, g_0).$$

By induction, we know that $\|\bar{e}(k)\|_\infty \leq K_2(h, \gamma, g_0)$ for all $k \geq 0$, which means that the quantizer will never be saturated if (9.58) is satisfied. By the definition of $w(k)$ and $0 < \gamma < 1$, (9.9) is obtained. The proof of (9.59) and (9.60) is analogous to that of (9.37) and (9.38) which is omitted here. $\square$

Remark 9.5. *Note that the result is more conservative for the time-varying delay case than that for the constant delay case. For the same consensus gain h and scaling parameter γ , we need a higher transmission bit rate for time-varying delays case. Nevertheless, K_2 can also approach $\frac{1}{2}$ when $h \rightarrow 0^+$ and $\gamma \rightarrow 1^-$, which implies the attainability of 1-bit information exchange while guaranteeing the average consensus.*

9.3 Numerical Examples

We consider a network with 5 agents and 0 – 1 weights over the edges. The edges of the graph are initially randomly generated with probability 0.5. By calculation, we have $\lambda_2 = 3$, $\lambda_N = 5$ and $D^* = 4$. The initial conditions of the agents are

$$x_1(0) = 0.8449, \quad x_2(0) = 0.7531, \quad x_3(0) = 0.7030, \quad x_4(0) = 0.2466, \quad x_5(0) = 0.0399.$$

We assume that all the transmitted data suffers from 1 step delay. We choose $h = 0.0005$. $\gamma = 0.9993$ and $g_0 = 2$. According to (9.33) we have $K_1 = 0.8539$. From Theorem 9.1 we know that one-bit quantizer can be used. The state trajectories are shown in Fig. 9.4. From the figure we see that the average value of the initial states, 0.5175, is achieved asymptotically.

Figure 9.4: State trajectories of 5 agents with $h = 0.0005$ and $\gamma = 0.9993$.

Following Remark 9.3, there is a trade-off in choosing the parameter h , γ and the data rate. From the above example we see that 1-bit rate communication requires that γ is close enough to 1. In fact, the bit number is a conservative estimate, and in practice, fewer bits may be enough when a smaller γ is chosen. If we take $h = 0.01$

and $\gamma = 0.99$, from (9.33) we get $K_1 = 19.31$. We still use one-bit quantizer, the state trajectories are shown in Fig. 9.5. We see that the average consensus is achieved asymptotically as well.

Figure 9.5: State trajectories of 5 agents with $h = 0.01$ and $\gamma = 0.99$.

9.4 Conclusion

We have considered the average consensus problem for first order discrete-time multi-agent systems with communication delays and limited data rate. The dynamical encoding and decoding with finite-level quantization has been adopted. It has been shown that when the network is connected, a distributed protocol can be designed such that the average consensus problem is solvable under a finite-level quantizer. In particular, no matter how large the delays are, we are always able to design a distributed protocol to ensure asymptotic average consensus with an exponential convergence rate using merely one bit information exchange between each pair of adjacent agents at each time step.

The encoder and decoder are node based, which makes the results restrictive. To extend the result to dynamically switching communication graphs, we need to design a channel based coding strategy, i.e. to design different encoder for different channel. Further, it is possible to consider random packet dropouts and/or random communication graphs when we adopt the channel based coding strategy. It is worthy mentioning that the receipt acknowledgement of each channel is required when the channel based coding strategy is used.

9.5 Appendix: Proof of Lemma 9.2

Before proving Lemma 9.1, we introduce the following Lemma.

Lemma 9.4. *Given a series of $n \times n$ matrices $F_1, F_2, \dots, F_l$ satisfying $\forall i = 1, \dots, l, \|F_i\| < 1$, we have the following result:*

$$\begin{aligned} & \|(I - F_1)^{-1} (I - F_2)^{-1} \dots (I - F_l)^{-1} - I\| \\ & \leq (1 - \|F_1\|)^{-1} (1 - \|F_2\|)^{-1} \dots (1 - \|F_l\|)^{-1} - 1, \end{aligned} \quad (9.73)$$

where $\|\cdot\|$ can be any induced norm.

Proof. $\|F_i\| < 1$ guarantees that $I - F_i$ is invertible. In fact, if $I - F_i$ is singular, we can find a vector v satisfying $(I - F_i)v = 0$. Thus

$$\|v\|_2 = \|F_i v\|_2 \leq \|F_i\| \|v\|_2 < \|v\|_2,$$

which is a contradiction. Therefore such v does not exist. Then for any $i = 1, 2, \dots, l$, we have

$$(I - F_i)^{-1} = \sum_{k=0}^{\infty} F_i^k,$$

and therefore

$$\|(I - F_1)^{-1} \dots (I - F_l)^{-1} - I\| = \left\| \left(\sum_{k=0}^{\infty} F_1^k \right) \left(\sum_{k=0}^{\infty} F_2^k \right) \dots \left(\sum_{k=0}^{\infty} F_l^k \right) - I \right\|$$

$$\begin{aligned}
&= \left\| \sum_{1 \leq i \leq l} F_i + \sum_{1 \leq i \leq j \leq l} F_i F_j + \cdots \right\| \\
&\leq \sum_{1 \leq i \leq l} \|F_i\| + \sum_{1 \leq i \leq j \leq l} \|F_i\| \|F_j\| + \cdots \\
&= (1 - \|F_1\|)^{-1} (1 - \|F_2\|)^{-1} \cdots (1 - \|F_l\|)^{-1} - 1,
\end{aligned}$$

which is (9.73). $\square$

Proof of Lemma 9.1:

It is straightforward that for any given θ satisfying (9.21), if $h < h_1$, then we have

$$\bar{\lambda}h + \rho_A(\bar{\lambda}, \theta)h^2 \leq \theta. \quad (9.74)$$

From (9.23) and (9.74) we see that $A(k)$ is invertible for $k = 0, 1, \dots, d-1$. Next, we shall show that (9.23) holds for all $k \geq 0$. Assume that $\|\Delta_A(k)\|_2 < \rho_A(\bar{\lambda}, \theta)h^2$ for $k = 0, \dots, r-1$, where $r \geq d$. According to (9.15)-(9.18), $\forall d \leq k \leq r$, we arrive at the following relationships:

$$\begin{aligned}
I - \gamma A(k) &= \sum_{i=0}^k hL_i(k) \gamma^{-i} (\Pi_{k-1, k-i}^A)^{-1} \\
&= \sum_{i=1}^k hL_i(k) \left[(I - (hL - \Delta_A(k-i)))^{-1} \cdots (I - (hL - \Delta_A(k-1)))^{-1} \right] \\
&\quad + hL_0(k) \\
&= hL - \Delta_A(k).
\end{aligned} \quad (9.75)$$

Denote $g(k) = \max_{j=1, \dots, d} \|hL - \Delta_A(k-j)\|_2$. It is clear that $\forall k \leq r$, we have $g(k) \leq \theta$. This is due to that $g(k) \leq \bar{\lambda}h + \rho_A(\bar{\lambda}, \theta)h^2$. According to (9.75), we have

$$\begin{aligned}
\|\Delta_A(r)\|_2 &= \left\| \sum_{i=1}^d hL_i(r) \left[\left(\Pi_{r-1, r-i}^{I-hL+\Delta_A} \right)^{-1} - I \right] \right\|_2 \\
&\leq h \|L(k)\|_2 \sum_{i=1}^d \left[(1 - g(r))^{-i} - 1 \right] \\
&\leq \bar{\lambda}h \frac{(1 - g(r))^{-d} - 1 - dg(r)}{g(r)}
\end{aligned}$$

$$\begin{aligned}
&\leq \bar{\lambda}h (\bar{\lambda}h + \rho_A(\bar{\lambda}, \theta)h^2) \frac{(1 - g(r))^{-d} - 1 - dg(r)}{g^2(r)} \\
&< h^2 (\bar{\lambda}^2 + \theta\rho_A(\bar{\lambda}, \theta)) \frac{(1 - \theta)^{-d} - 1 - d\theta}{\theta^2}.
\end{aligned} \tag{9.76}$$

The first inequality comes from Lemma A.2, Lemma 9.4 and the definition of $g(k)$. The third inequality is due to the fact $g(r) \leq \bar{\lambda}h + \rho_A(\bar{\lambda}, \theta)h^2$. The last inequality is obtained from the inequality $\bar{\lambda}h < \bar{\lambda}h + \rho_A h^2 \leq \theta$. According to the definition of ρ_A and inequality (9.76) we see that $\|\Delta_A(r)\|_2 < \rho_A(\bar{\lambda}, \theta)h^2$. It is clear that $A(r)$ is also invertible. By induction, $\|\Delta_A(k)\|_2 < \rho_A(\bar{\lambda}, \theta)h^2$ and $A(k)$ is invertible for all $k \geq 0$.

Next, we shall prove (9.26). According to (9.15)-(9.18) we have $\forall j = 1, 2, \dots, d$,

$$B_j(k) = \sum_{i=j}^d hL_i(k) \gamma^{-i-1} (\Pi_{k-j, k-i}^A)^{-1}. \tag{9.77}$$

Since $(1 - \theta)^l < \gamma < 1$, we know that $\forall j < d$, $\gamma^{-j} < \gamma^{-d} < (1 - \theta)^{-ld}$. Then, by comparing (9.22) and (9.77), we have

$$\begin{aligned}
\|\Delta_{B_j}(k)\|_2 &= \left\| \gamma^{-j} \sum_{i=j}^d hL_i(k) \left(\gamma^{j-i-1} (\Pi_{k-j, k-i}^A)^{-1} - I \right) \right\|_2 \\
&\leq h\gamma^{-j} \|L(k)\|_2 \sum_{i=j}^d \left[(1 - g(k))^{-(i-j+1)} - 1 \right] \\
&< \bar{\lambda}h\gamma^{-d} \frac{(1 - g(k))^{-d} - 1 - dg(k)}{g(k)} \\
&\leq \bar{\lambda}h\gamma^{-d} (\bar{\lambda}h + \rho_A(\bar{\lambda}, \theta)h^2) \frac{(1 - g(k))^{-d} - 1 - dg(k)}{g^2(k)} \\
&< h^2 (\bar{\lambda}^2 + \theta\rho_A(\bar{\lambda}, \theta)) \frac{(1 - \theta)^{-d} - 1 - d\theta}{\theta^2(1 - \theta)^{ld}}.
\end{aligned}$$

It is known that for any given $n \times n$ matrix B , $\|B\|_\infty \leq \sqrt{n} \|B\|_2$, then we have

$$\|\Delta_{B_j}(k)\|_\infty \leq \sqrt{N} \|\Delta_{B_j}(k)\|_2 < \rho_B(\bar{\lambda}, \theta, l)h^2.$$

□

Chapter 10

Conclusions and Future Work

10.1 Conclusions

In this thesis, we have discussed some issues of controller design for single agent and multi-agent time-delay systems. For single agent systems, linear quadratic regulation and tracking was investigated for systems with multiple delays in a single input channel. For multi-agent systems, some complex distributed consensus control problems of MASs with communication and input delays were discussed. We summarize the contributions of the thesis in the following aspects.

We have investigated the LQR problem for both linear discrete-time systems and linear continuous-time systems with multiple delays in a single input channel. By introducing an auxiliary backward stochastic system and through re-organization of output measurements, the deterministic quadratic optimization problem was first converted into a smoothing estimation problem. A new controller derivation method was developed which overcomes the difficulty posed by multiple delays in a single input channel. Our solution is given in terms of the solution of a two-dimensional Riccati difference equation for the discrete-time case and a Riccati partial differential equation for the continuous-time case.

The infinite horizon LQR problem has also been solved and conditions for convergence and stability have been provided. Our approach applies projection approach

as opposed to the complicated infinite-dimensional systems theory approach for continuous-time systems and offers significant computational advantages as compared to the state augmentation approach for discrete-time systems. The regulation results were extended to linear quadratic tracking problem. An auxiliary backward deterministic system was introduced for the explicit expression of the optimal controller.

We then looked into consensus control of multi-agent systems with integrator node dynamics. Firstly, we comparatively studied two control protocols for a leader-follower problem, a standard local voting protocol, and a protocol that includes information about the control inputs of neighbors. We showed that the latter one leads to a better performance. However, we noticed its implementation problem and addressed it by introducing a small delay for neighbors' control input related term. The consensusability and convergence rate were analyzed for the latter protocol with time delays. This protocol was successfully implemented to UGV formation.

Secondly, we considered MASs with both time delays and noises in transmission channels. By introducing a decaying consensus gain, the consensus of the whole system can be achieved in a stochastic sense and necessary and sufficient conditions have been given for strong consensus and mean square consensus when communication protocol is static. We also developed sufficient conditions for dynamically switching protocols and randomly switching protocols cases.

Last, we studied the average consensus problem for MASs with time delays and data rate constraint. A distributed consensus protocol was proposed based on dynamic encoding and decoding. We have shown that for a connected network, as long as the time delays are bounded, the average consensus can be achieved with finite communication data rate. Moreover, merely one bit information exchange between each pair of adjacent agents at each time step suffices to guarantee the average consensus.

10.2 Future Work

1. In Part I, we have established a systematical approach for optimal control of linear systems with multiple delays in a single input channel. Based on this fundamental result, some more complicated problems can be considered. For example, linear quadratic Gaussian (LQG) and H_∞ controllers are much more useful for practical systems. Most of the previous works focus on the LQG and H_∞ control problems for single input delay systems. [77] and [58] investigated the control problem for systems with delay in disturbance and control inputs, containing the H_∞ control with preview as a special case. [135] considered the H_∞ control problem for linear systems of multiple input channel, with single delay in each channel. It would be interesting to extend the results in Part I to LQG, H_∞ control, and mixed H_2/H_∞ control problems for systems with multiple delays in single input channel.
2. Robustness analysis should be brought into consideration for single agent input delay systems. For example, delay is not always exactly known, especially for stochastic delay systems, so delay mismatch is unavoidable. For another example, there is existing literature on control problem for systems with multiplicative noise. However, there exist few works on input delay systems with multiplicative noise. [110] provides a way to solve the stochastic LQR problem for single input delay systems with state and input dependent noise by exploiting the dynamic programming approach. The control problem for multiple input delay systems with multiplicative noise is still challenging due to the interactions between the delayed input channels and the stochastic parameter uncertainties.
3. In Chapter 8, we have investigated consensus problem for MASs with first order discrete-time system. The result can be extended to agents with continuous-time dynamics. On the other hand, we have considered additive noises. However, sometimes the transmission channel has multiplicative noises, e.g. fading

channels. In the future, we can consider consensus with multiplicative noises and delays in transmission channel. The convergence rate should also be studied.

4. In Part II, we have investigated some complex consensus problems with communication and input delays. However, the MASs are with first order linear dynamics. It is expected that the obtained results can be extended to the agents with more general linear systems or even nonlinear systems.
5. In Chapter 7, we have successfully implemented the result to UGV formation control. However, in the real systems, transmission noises and data rate constraint are unavoidable. It is expected that the algorithms derived in Chapter 8 and 9 can be implemented in UGV formation and even UAV formation control.

Appendix A

Preliminaries in Graph Theory

A directed graph, denoted by $\mathcal{G}$, with vertex set $\mathcal{V}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is often used to model communications among agents. For convenience, we define $\mathcal{V} = \{1, 2, \dots, n\}$, where n is the number of vertices. We consider that $(i, j) \in \mathcal{E}$ if and only if vertex (node) i can send its information to vertex j . If $(i, i) \in \mathcal{E}$, we say that vertex i has self-loop. In this thesis, we do not consider vertices with self-loop. The set of neighbors of vertex i is denoted by $\mathcal{N}_i = \{j \mid j \in \mathcal{V}, (j, i) \in \mathcal{E}\}$. We denote by $a_{i,j} \geq 0$ the weighting on the edge (j, i) . When $j \notin \mathcal{N}_i$, which means there is no information flow from vertex j to vertex i , $a_{i,j} = 0$, otherwise, $a_{i,j} > 0$. For any pair of vertices (i, j) , if $a_{i,j} = a_{j,i}$, the graph is called an undirected graph.

Usually, we use the triplet $\{\mathcal{V}, \mathcal{E}, A\}$ to describe a graph $\mathcal{G}$, where $A \in \mathbb{R}^{n \times n}$ is the adjacency matrix associated with $\mathcal{G}$ and $\forall i, j \in \mathcal{V}$, $A^{i,j} = a_{i,j}$. There is a path from vertex i to vertex j if there exists a sequence $l_1, \dots, l_p \in \mathcal{V}$ satisfying $(i, l_1), (l_1, l_2), \dots, (l_p, j) \in \mathcal{E}$ where $i, j, l_1, \dots, l_p$ are distinct vertices. Given a graph $\mathcal{G}$, it contains a spanning tree if there exists at least one vertex i such that for any other vertex j , there is a path from i to j . An undirected graph is connected if it contains a spanning tree.

We call $d_i \triangleq \sum_{j \in \mathcal{V}} a_{i,j}$ the in-degree of vertex i and $D \triangleq \text{diag}\{d_1, \dots, d_n\}$ the

in-degree matrix of $\mathcal{G}$. The Laplacian matrix L of the graph $\mathcal{G}$ is defined as follows:

$$L \triangleq D - A, \quad L^{i,j} = \begin{cases} -a_{i,j}, & i \neq j, \\ d_i, & i = j. \end{cases}$$

By denoting all the eigenvalues of L as λ_i , $i = 1, 2, \dots, n$, some properties of the Laplacian matrix are recalled below [98]:

Lemma A.1. *For a directed graph $\mathcal{G}$, the eigenvalues of the Laplacian matrix of $\mathcal{G}$ are λ_i , $i = 1, 2, \dots, n$. Suppose that $\forall i < j$, $\text{Re}(\lambda_i) \leq \text{Re}(\lambda_j)$, we have the following properties:*

1. $\lambda_1 = 0$.
2. $\text{Re}(\lambda_2) > 0$ if and only if the graph $\mathcal{G}$ contains a directed spanning tree.

Furthermore, if the graph $\mathcal{G}$ is an undirected graph, its Laplacian matrix $L(\mathcal{G})$ is symmetric and all the eigenvalues of $L(\mathcal{G})$ are real. If the undirected graph is connected, then $\lambda_2 > 0$ and for a vector δ ,

$$\min_{\delta \neq 0, \mathbf{1}'_{|\mathcal{V}|} \delta = 0} \frac{\delta' L \delta}{\|\delta\|_2} = \lambda_2.$$

Lemma A.2. *For Laplacian matrices L_1 and L_2 corresponding to undirected graphs satisfying $L_1 \leq L_2$, we have $\|L_1\|_2 \leq \|L_2\|_2$ and $\|L_1\|_\infty \leq \|L_2\|_\infty$.*

Proof. It follows directly from the properties of Laplacian matrix [36] and the details are omitted. $\square$

A graph is balanced if the in-degree d_i and the out-degree $\tilde{d}_i \triangleq \sum_{j \in \mathcal{V}} a_{j,i}$ are equal for all $i \in \mathcal{V}$. For example, an undirected graph is a kind of balanced graph.

At last, the quotient graph should be introduced [13], which is often used for analyzing the consensus problem for MASs with communication delays. The quotient graph, denoted by $\mathcal{G}_Q$ is defined as the communication graph by ignoring delays. In other words, there is an edge $(i, j) \in \mathcal{E}(\mathcal{G}_Q)$ if (delayed) information from agent i is successfully received by agent j .

Author's Publications

Journal papers:

1. S. Liu, L. Xie, and H. Zhang, "Distributed consensus for multi-agent systems with delays and noises in transmission channels," *Automatica*, vol. 47, no. 5, pp. 920-934, 2011 (regular paper).
2. S. Liu, L. Xie, and F. Lewis, "Synchronization of Multi-Agent Systems with Delayed Input Information from Neighbors," *Automatica*, vol. 47, no. 10, pp. 2152-2164, 2011 (regular paper).
3. S. Liu, T. Li, and L. Xie, "Distributed Consensus for Multi-agent Systems with Communication Delays and Limited Data Rate," *accepted by SIAM Journal on Control and Optimization*, 2011.
4. S. Liu, L. Xie, and H. Zhang, "Linear quadratic tracking problem for discrete-time systems with multiple delays in single input channel," *International Journal of Robust and Nonlinear Control*, vol. 20, no 12, pp. 1379-1394, 2010.
5. S. Liu, L. Xie, and H. Zhang, "Infinite Horizon LQR for Systems with Multiple Delays in a Single Input Channel," *Journal of Control Theory and Applications*, vol. 8, no. 3, pp. 368-374, 2010.

Conference papers:

1. S. Liu, L. Xie, and H. Zhang, "Linear Quadratic Regulation for Discrete-time Systems with Multiple Delays in Single Input Channel," in *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, 2008, pp. 13336-13341.
2. S. Liu, L. Xie, and H. Zhang, "Linear quadratic regulation for linear continuous-time systems with distributed input delay," in *Proceedings of the 27th Chinese Control Conference*, Kunming, 2008, pp. 195-199.
3. H. Wang, H. Zhang, L. Xie, and S. Liu, "Discrete time H_∞ preview control," in *Proceedings of the 7th Asian Control Conference*, Hong Kong, 2009, pp. 57-62.
4. S. Liu, L. Xie, and T. Li, "Distributed consensus for multi-agent systems with communication delays and limited data rate," in *Proceedings of the 8th World Congress on Intelligent Control and Automation*, Jinan, 2010, pp. 904-909.
5. S. Liu and L. Xie, "Consensus for discrete multi-agent systems with bounded delays under static and randomly switching topologies," in *Proceedings of the 29th Chinese Control Conference*, Beijing, 2010, pp. 5809-5814.
6. S. Liu, L. Xie, and F. L. Lewis, "Synchronization of multi-agent systems with delayed input information from neighbors," in *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, 2010, pp. 4529-4534.
7. S. Liu, L. Xie, and H. Zhang, "Consensus protocols for multi-agent systems with delays and noises in transmission channels," in *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, 2010, pp. 5857-5862.
8. S. Liu, C. Chen, L. Xie, and Y.-H. Chang, "Formation control of multi-robot systems," in *Proceedings of the 11th International Conference on Control Automation Robotics & Vision*, Singapore, 2010, pp. 1057-1062.

Bibliography

- [1] Y. Alekal, P. Brunovsky, D. Chyung, and E. Lee, “The quadratic problem for systems with time delays,” *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 673–687, 1971.
- [2] B. Anderson and J. Moore, *Optimal Filtering*. New Jersey: Prentice-Hall Information and System Sciences Series, 1979.
- [3] D. Angeli and P.-A. Bliman, “Stability of leaderless discrete-time multi-agent systems,” *Mathematics of Control, Signals, and Systems*, vol. 18, no. 4, pp. 293–322, 2006.
- [4] —, “Tight estimates for convergence of some non-stationary consensus algorithms,” *Systems & Control Letters*, vol. 57, no. 12, pp. 996–1004, 2008.
- [5] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, “Broadcast gossip algorithms for consensus,” *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2748–2761, 2009.
- [6] R. B. Bapat and T. E. S. Raghavan, *Nonnegative matrices and applications*. Cambridge University Press, 1997.
- [7] D. Bauso, L. Giarr, and R. Pesenti, “Non-linear protocols for optimal distributed consensus in networks of dynamic agents,” *Systems & Control Letters*, vol. 55, no. 11, pp. 918–928, 2006.

- [8] A. Bensoussan, G. Da Prato, M. Delfour, and S. Mitter, *Representation and Control of Infinite Dimensional Systems*. Verlag Basel, Switzerland: Birkhäuser, 2007.
- [9] S. Bittanti, A. Laub, and J. Willems, *The Riccati Equations*. Berlin Heidelberg, New York: Springer-Verlag, 1991.
- [10] P.-A. Bliman and G. Ferrari-Trecate, “Average consensus problems in networks of agents with delayed communications,” *Automatica*, vol. 44, no. 8, pp. 1985–1995, 2008.
- [11] A. Callender, D. R. Hartree, and A. Porter, “Time-lag in a control system,” *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 235, no. 756, pp. 415–444, 1936.
- [12] M. Cao, A. S. Morse, and B. D. O. Anderson, “Reaching a consensus in a dynamically changing environment: A graphical approach,” *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 575–600, 2008.
- [13] —, “Reaching a consensus in a dynamically changing environment: Convergence rates, measurement delays, and asynchronous events,” *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 601–623, 2008.
- [14] Y. Cao and W. Ren, “Sampled-data discrete-time coordination algorithms for double-integrator dynamics under dynamic directed interaction,” *International Journal of Control*, vol. 83, no. 3, pp. 506 – 515, 2010.
- [15] R. Carli and F. Bullo, “Quantized coordination algorithms for rendezvous and deployment,” *SIAM Journal on Control and Optimization*, vol. 48, no. 3, pp. 1251–1274, 2009.
- [16] R. Carli, F. Bullo, and S. Zampieri, “Quantized average consensus via dynamic coding/decoding schemes,” *International Journal of Robust and Non-linear Control*, vol. 20, no. 2, pp. 156–175, 2010.

- [17] R. Carli, F. Fagnani, P. Frasca, and S. Zampieri, “Gossip consensus algorithms via quantized communication,” *Automatica*, vol. 46, no. 1, pp. 70–80, 2010.
- [18] S. Chan, G. Goodwin, and K. Sin, “Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems,” *IEEE Transactions on Automatic Control*, vol. 29, no. 2, pp. 110–118, 1984.
- [19] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*. Springer Verlag, 2003.
- [20] J. Chu, “Application of a discrete optimal tracking controller to an industrial electric heater with pure delays,” *Journal of Process Control*, vol. 5, no. 1, pp. 3–8, 1995.
- [21] D. Chyung, “Discrete systems with delays in control,” *IEEE Transactions on Automatic Control*, vol. 14, no. 2, pp. 196–197, 1969.
- [22] R. Datko, J. Lagnese, and M. P. Polis, “An example on the effect of time delays in boundary feedback stabilization of wave equations,” *SIAM Journal on Control and Optimization*, vol. 24, no. 1, pp. 152–156, 1986.
- [23] M. C. Delfour, “The linear-quadratic optimal control problem with delays in state and control variables: A state space approach,” *SIAM Journal on Control and Optimization*, vol. 24, no. 5, pp. 835–883, 1986.
- [24] Z. Deng, *Kalman Filter and Wiener Filter: Modern Time Sequences Analysis Method [M]*. Harbin: Harbin Polytechnique University Press, 2001.
- [25] S. Di Cairano, A. Pasini, A. Bemporad, and R. M. Murray, “Convergence properties of dynamic agents consensus networks with broken links,” in *Proceedings of American Control Conference*, Seattle, 2008, pp. 1362–1367.
- [26] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, “Gossip algorithms for distributed signal processing,” *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

- [27] D. V. Dimarogonas and K. J. Kyriakopoulos, “On the rendezvous problem for multiple nonholonomic agents,” *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 916–922, 2007.
- [28] M. G. Earl and S. H. Strogatz, “Synchronization in oscillator networks with delayed coupling: A stability criterion,” *Physical Review E*, vol. 67, no. 3, p. 036204, 2003.
- [29] D. Eller, J. Aggarwal, and H. Banks, “Optimal control of linear time-delay systems,” *IEEE Transactions on Automatic Control*, vol. 14, no. 6, pp. 678–687, 1969.
- [30] F. Fagnani and S. Zampieri, “Randomized consensus algorithms over large scale networks,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 634–649, 2008.
- [31] —, “Average consensus with packet drop communication,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 102–133, 2009.
- [32] J. A. Fax and R. M. Murray, “Information flow and cooperative control of vehicle formations,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465–1476, 2004.
- [33] Y. Fiagbedzi and A. Pearson, “Feedback stabilization of linear autonomous time lag systems,” *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 847–855, 1986.
- [34] E. Fridman and U. Shaked, “Special issue on time-delay systems,” *International Journal of Robust and Nonlinear Control*, vol. 13, no. 9, pp. 791–792, 2003.
- [35] T. Fujinaka and M. Araki, “Discrete-time optimal control of systems with unilateral time-delays,” *Automatica*, vol. 23, no. 6, pp. 763–765, 1987.

- [36] C. D. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer, 2001.
- [37] J. R. Grad and K. A. Morris, “Solving the linear quadratic optimal control problem for infinite-dimensional systems,” *Computers & Mathematics with Applications*, vol. 32, no. 9, pp. 99–119, 1996.
- [38] S. Graham and P. R. Kumar, “Time in general-purpose control systems: the control time protocol and an experimental evaluation,” in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 4, Nassau, 2004, pp. 4004–4009.
- [39] K. Gu, V. Kharitonov, and J. Chen, *Stability of Time-Delay Systems*. Birkhäuser, 2003.
- [40] L. Guo, “Time-varying stochastic systems: Stability, estimation and control,” *Jilin Science and Technology Press, Jilin, China*, 1990.
- [41] D. R. Hartree, A. Porter, A. Callender, and A. B. Stevenson, “Time-lag in a control system. II,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 161, no. 907, pp. 460–476, 1937.
- [42] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control: A Unified Approach to H_2 and H_∞ Theories*. SIAM, 1999.
- [43] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [44] M. Huang, S. Dey, G. N. Nair, and J. H. Manton, “Stochastic consensus over noisy networks with markovian and arbitrary switches,” *Automatica*, vol. 46, no. 10, pp. 1571–1583, 2010.
- [45] M. Huang and J. H. Manton, “Stochastic approximation for consensus seeking: Mean square and almost sure convergence,” in *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, 2007, pp. 306–311.

- [46] ———, “Stochastic lyapunov analysis for consensus algorithms with noisy measurements,” in *Proceedings of American Control Conference*, New York, 2007, pp. 1419–1424.
- [47] ———, “Stochastic consensus seeking with measurement noise: Convergence and asymptotic normality,” in *Proceedings of the American Control Conference*, Seattle, 2008, pp. 1337–1342.
- [48] ———, “Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 134–161, 2009.
- [49] A. Ichikawa, “Quadratic control of evolution equations with delays in control,” *SIAM Journal on Control and Optimization*, vol. 20, no. 5, pp. 645–668, 1982.
- [50] K. Ito and H. T. Tran, “Linear quadratic regulator problem for infinite dimensional linear systems with delays in control,” in *Proceedings of the 27th IEEE Conference on Decision and Control*, H. T. Tran, Ed., 1988, pp. 2012–2017.
- [51] A. Jadbabaie, L. Jie, 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, 2003.
- [52] M. Jamshidi and M. Heidari, “Application of dynamic programming to control khuzestan water resources system,” *Automatica*, vol. 13, no. 3, pp. 287–293, 1977.
- [53] S. Kar and J. M. F. Moura, “Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, 2009.
- [54] S. M. Karbassi, “Time-optimal control of disturbance-rejection tracking systems for discrete-time time-delayed systems by state feedback,” *Computers & Mathematics with Applications*, vol. 50, no. 8-9, pp. 1415–1424, 2005.

- [55] A. Kashyap, T. Basar, and R. Srikant, “Quantized consensus,” *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.
- [56] B. Keulen, *H_∞ control for Distributed Parameter Systems: A State Space Approach*. Birkhäuser, 1993.
- [57] S. A. Khan and A. P. Engelbrecht, “A new fuzzy operator and its application to topology design of distributed local area networks,” *Information Sciences*, vol. 177, no. 13, pp. 2692–2711, 2007.
- [58] A. Kojima and S. Ishijima, “Formulas on preview and delayed H^∞ control,” *IEEE Transactions on Automatic Control*, vol. 51, no. 12, pp. 1920–1937, 2006.
- [59] E. Kokiopoulou and P. Frossard, “Polynomial filtering for fast convergence in distributed consensus,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 342–354, 2009.
- [60] V. B. Kolmanovskii and A. Myshkis, *Applied Theory of Functional Differential Equations*. Kluwer Academic Publishers, 1992.
- [61] H. Kwakernaak, “Optimal filtering in linear systems with time delays,” *IEEE Transactions on Automatic Control*, vol. 12, no. 2, pp. 169–173, 1967.
- [62] W. H. Kwon and S. Han, *Receding Horizon Control: Model Predictive Control for State Models*. Springer, 2005.
- [63] G. Lafferriere, A. Williams, J. Caughman, and J. J. P. Veerman, “Decentralized control of vehicle formations,” *Systems & Control Letters*, vol. 54, no. 9, pp. 899–910, 2005.
- [64] T. Li, M. Fu, L. Xie, and J. F. Zhang, “Distributed consensus with limited communication data rate,” *provisionally accepted by IEEE Transaction on Automatic Control*, 2009.

- [65] 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, 2010.
- [66] 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.
- [67] F. L. Lian, J. Moyne, and D. Tilbury, “Analysis and modeling of networked control systems: MIMO case with multiple time delays,” pp. 4306–4312, 2001.
- [68] J. Lin, A. S. Morse, and B. D. O. Anderson, “The multi-agent rendezvous problem,” in *Proceedings of the 42nd IEEE Conference on Decision and Control*, vol. 2, 2003, pp. 1508–1513 Vol.2.
- [69] P. Lin and Y. Jia, “Consensus of second-order discrete-time multi-agent systems with nonuniform time-delays and dynamically changing topologies,” *Automatica*, vol. 45, no. 9, pp. 2154–2158, 2009.
- [70] Z. Lin, B. Francis, and M. Maggiore, “State agreement for continuous-time coupled nonlinear systems,” *SIAM Journal on Control and Optimization*, vol. 46, no. 1, pp. 288–307, 2007.
- [71] M. Liu, H. Zhang, G. Duan, and B. Xu, “Tracking control with multiple input delays for continuous-time systems,” in *Proceedings of the 6th World Congress on Intelligent Control and Automation*, vol. 1, Da Lian, 2006, pp. 290–294.
- [72] S. Liu, L. Xie, and H. Zhang, “Linear quadratic regulation for discrete-time systems with multiple delays in single input channel,” in *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, 2008, pp. 13 336–13 341.
- [73] X. Liu, W. Lu, and T. Chen, “Consensus of multi-agent systems with unbounded time-varying delays,” *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2396–2401, 2010.

- [74] J. Lu, D. W. C. Ho, J. Kurths, uuml, and rgen, “Consensus over directed static networks with arbitrary finite communication delays,” *Physical Review E*, vol. 80, no. 6, p. 066121, 2009.
- [75] C. Ma, T. Li, and J. Zhang, “Consensus control for leader-following multi-agent systems with measurement noises,” *Journal of Systems Science and Complexity*, vol. 23, no. 1, pp. 35–49, 2010.
- [76] A. Manitius and A. Olbrot, “Finite spectrum assignment problem for systems with delays,” *IEEE Transactions on Automatic Control*, vol. 24, no. 4, pp. 541–552, 1979.
- [77] G. Meinsma and L. Mirkin, “ H_∞ control of systems with multiple I/O delays via decomposition to adobe problems,” *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 199–211, 2005.
- [78] P. P. Menon and C. Edwards, “Decentralised static output feedback stabilisation and synchronisation of networks,” *Automatica*, vol. 45, no. 12, pp. 2910–2916, 2009.
- [79] A. A. Moelja and G. Meinsma, “ H_2 -optimal control of systems with multiple I/O delays: Time domain approach,” *Automatica*, vol. 41, no. 7, pp. 1229–1238, 2005.
- [80] L. Moreau, “Stability of multiagent systems with time-dependent communication links,” *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 169–182, 2005.
- [81] U. Munz, A. Papachristodoulou, and F. Allgover, “Delay robustness in consensus problems,” *Automatica*, vol. 46, no. 8, pp. 1252–1265, 2010.
- [82] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, “On distributed averaging algorithms and quantization effects,” *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2506–2517, 2009.

- [83] A. Nedic and A. Ozdaglar, “Convergence rate for consensus with delays,” *Journal of Global Optimization*, vol. 47, no. 3, pp. 437–456, 2010.
- [84] R. Olfati-Saber, “Distributed kalman filtering for sensor networks,” in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.
- [85] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
- [86] 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, 2004.
- [87] R. Olfati-Saber and J. S. Shamma, “Consensus filters for sensor networks and distributed sensor fusion,” in *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference*, 2005, pp. 6698–6703.
- [88] A. Olshevsky and J. N. Tsitsiklis, “Convergence speed in distributed consensus and averaging,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.
- [89] A. Papachristodoulou, A. Jadbabaie, and U. Mnz, “Effects of delay in multi-agent consensus and oscillator synchronization,” *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1471–1477, 2010.
- [90] S. S. Pereira and A. Pages-Zamora, “Mean square convergence of consensus algorithms in random WSNs,” *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2866–2874, 2010.
- [91] R. Pindyck, “An application of the linear quadratic tracking problem to economic stabilization policy,” *IEEE Transactions on Automatic Control*, vol. 17, no. 3, pp. 287–300, 1972.

- [92] ———, “The discrete-time tracking problem with a time delay in the control,” *IEEE Transactions on Automatic Control*, vol. 17, no. 3, pp. 397–398, 1972.
- [93] J. B. Pomet, B. Thuilot, G. Bastin, and G. Campion, “A hybrid strategy for the feedback stabilization of nonholonomic mobile robots,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, 1992, pp. 129–134.
- [94] A. J. Pritchard and D. Salamon, “The linear quadratic control problem for infinite dimensional systems with unbounded input and output operators,” *SIAM Journal on Control and Optimization*, vol. 25, no. 1, pp. 121–144, 1987.
- [95] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer Verlag, 2009.
- [96] Z. Qu, J. Wang, and R. A. Hull, “Cooperative control of dynamical systems with application to autonomous vehicles,” *IEEE Transactions on Automatic Control*, vol. 53, no. 4, pp. 894–911, 2008.
- [97] W. Ren, “Consensus strategies for cooperative control of vehicle formations,” *Control Theory & Applications, IET*, vol. 1, no. 2, pp. 505–512, 2007.
- [98] W. Ren and R. W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*, ser. Communications and Control Engineering. London: Springer, 2008.
- [99] W. Ren, “Multi-vehicle consensus with a time-varying reference state,” *Systems & Control Letters*, vol. 56, no. 7-8, pp. 474–483, 2007.
- [100] ———, “Consensus tracking under directed interaction topologies: Algorithms and experiments,” *IEEE Transactions on Control Systems Technology*, vol. 18, no. 1, pp. 230–237, 2010.
- [101] 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.

- [102] K. R. Santarelli and M. A. Dahleh, “Optimal controller synthesis for a class of lti systems via switched feedback,” *Systems & Control Letters*, vol. 59, no. 3-4, pp. 258–264, 2010.
- [103] E. Seneta, *Non-Negative Matrices and Markov Chains*. Springer, 2006.
- [104] J. S. Shamma and J. Shamma, *Cooperative Control of Distributed Multi-Agent Systems*. Wiley-Interscience, 2008.
- [105] M. G. Singh and H. Tamura, “Modelling and hierarchical optimization for oversaturated urban road traffic networks,” *International Journal of Control*, vol. 20, no. 6, pp. 913–934, 1974.
- [106] O. J. M. Smith, “Closer control of loops with dead time,” *Chemical Engineering Progress*, vol. 53, no. 5, pp. 217–219, 1957.
- [107] —, *Feedback Control Systems*. New York: McGraw-Hill Education, 1958.
- [108] R. Smith and F. Hadaegh, “Closed-loop dynamics of cooperative vehicle formations with parallel estimators and communication,” *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1404–1414, 2007.
- [109] S. L. Smith, M. E. Broucke, and B. A. Francis, “Curve shortening and the rendezvous problem for mobile autonomous robots,” *IEEE Transactions on Automatic Control*, vol. 52, no. 6, pp. 1154–1159, 2007.
- [110] X. Song, H. Zhang, and L. Xie, “Stochastic linear quadratic regulation for discrete-time linear systems with input delay,” *Automatica*, vol. 45, no. 9, pp. 2067–2073, 2009.
- [111] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, “Dynamic consensus on mobile networks,” in *Proceedings of the 16th IFAC world congress*, Prague, Czech, 2005.
- [112] G. Tadmor, “The standard H_∞ problem in systems with a single input delay,” *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 382–397, 2000.

- [113] G. Tadmor and L. Mirkin, “ H_∞ control and estimation with preview-part II: fixed-size ARE solutions in discrete time,” *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 29–40, 2005.
- [114] G. Tadmor, “ H_∞ interpolation in systems with commensurate input lags,” *SIAM Journal on Control and Optimization*, vol. 27, no. 3, pp. 511–526, 1989.
- [115] A. Tahbaz-Salehi and A. Jadbabaie, “A necessary and sufficient condition for consensus over random networks,” *IEEE Transactions on Automatic Control*, vol. 53, no. 3, pp. 791–795, 2008.
- [116] H. Tamura, “Decentralized optimization for distributed-lag models of discrete systems,” *Automatica*, vol. 11, no. 6, pp. 593–602, 1975.
- [117] H. Tan, S. Shu, and F. Lin, “An optimal control approach to robust tracking of linear systems,” *International Journal of Control*, vol. 82, no. 3, pp. 525–540, 2009.
- [118] G. Tang, H. Sun, and H. Pang, “Approximately optimal tracking control for discrete time-delay systems with disturbances,” *Progress in Natural Science*, vol. 18, no. 2, pp. 225–231, 2008.
- [119] Y.-P. Tian and C.-L. Liu, “Consensus of multi-agent systems with diverse input and communication delays,” *IEEE Transactions on Automatic Control*, vol. 53, no. 9, pp. 2122–2128, 2008.
- [120] S. E. Tuna, “Growth rate of switched homogeneous systems,” *Automatica*, vol. 44, no. 11, pp. 2857–2862, 2008.
- [121] K. Uchida, E. Shimemura, T. Kubo, and N. Abe, “The linear-quadratic optimal control approach to feedback control design for systems with delay,” *Automatica*, vol. 24, no. 6, pp. 773–780, 1988.
- [122] R. B. Vinter and R. H. Kwong, “The infinite time quadratic control problem for linear systems with state and control delays: An evolution equation

- approach,” *SIAM Journal on Control and Optimization*, vol. 19, no. 1, pp. 139–153, 1981.
- [123] H.-H. Wang, “Optimal tracking for discrete-time systems with input delays,” in *Proceedings of the Chinese Control and Decision Conference*, Yan Tai, 2008, pp. 4033–4037.
- [124] Q. G. Wang, T. H. Lee, and K. K. Tan, *Finite Spectrum Assignment for Time-delay Systems*. Springer, 1998.
- [125] P. Wei, Q. Guan, W. Yu, and L. Wang, “Easily testable necessary and sufficient algebraic criteria for delay-independent stability of a class of neutral differential systems,” *Systems & Control Letters*, vol. 57, no. 2, pp. 165–174, 2008.
- [126] D. Williams, *Probability with Martingales*. Cambridge University Press, 1991.
- [127] J. Wolfowitz, “Products of indecomposable, aperiodic, stochastic matrices,” *Proceedings of the American Mathematical Society*, vol. 14, no. 5, pp. 733–737, 1963.
- [128] F. Xiao and L. Wang, “Consensus protocols for discrete-time multi-agent systems with time-varying delays,” *Automatica*, vol. 44, no. 10, pp. 2577–2582, 2008.
- [129] F. Xiao, L. Wang, J. Chen, and Y. Gao, “Finite-time formation control for multi-agent systems,” *Automatica*, vol. 45, no. 11, pp. 2605–2611, 2009.
- [130] L. Xiao, S. Boyd, and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” in *4th International Symposium on Information Processing in Sensor Networks*, 2005, pp. 63–70.
- [131] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

- [132] L. Xiao, S. Boyd, and S.-J. Kim, “Distributed average consensus with least-mean-square deviation,” *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.
- [133] 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.
- [134] H. Zhang, G. Duan, and L. Xie, “Linear quadratic regulation for linear time-varying systems with multiple input delays,” *Automatica*, vol. 42, no. 9, pp. 1465–1476, 2006.
- [135] H. Zhang and L. Xie, *Control and Estimation of Systems with Input/Output Delays*. Berlin: Springer, 2007.
- [136] ———, “Linear quadratic regulation for systems with time-varying delay,” in *Proceedings of the 46th IEEE Conference on Decision and Control*, X. Lihua, Ed., New Orleans, 2007, pp. 2974–2979.
- [137] H. Zhang, L. Xie, and Y. C. Soh, “A unified approach to linear estimation for discrete-time systems. I: H_2 estimation,” in *Proceedings of the 40th IEEE Conference on Decision and Control*, X. Lihua, Ed., vol. 3, 2001, pp. 2917–2922.
- [138] ———, “A unified approach to linear estimation for discrete-time systems. II: H_∞ estimation,” in *Proceedings of the 40th IEEE Conference on Decision and Control*, X. Lihua, Ed., vol. 3, 2001, pp. 2923–2928.
- [139] W.-A. Zhang and L. Yu, “Modelling and control of networked control systems with both network-induced delay and packet-dropout,” *Automatica*, vol. 44, no. 12, pp. 3206–3210, 2008.
- [140] Q. C. Zhong, “On standard H_∞ control of processes with a single delay,” *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 1097–1103, 2003.

- [141] J. Zhou and Q. Wang, “Convergence speed in distributed consensus over dynamically switching random networks,” *Automatica*, vol. 45, no. 6, pp. 1455–1461, 2009.
- [142] K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice-Hall Englewood Cliffs, NJ, 1998.