CONTROL AND ESTIMATION WITH LIMITED INFORMATION: A GAME-THEORETIC APPROACH

BY

JUN MOON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Tamer Başar, Chair
Assistant Professor Mohamed Ali Belabbas
Professor Geir E. Dullerud
Professor Bruce Hajek

ABSTRACT

Modern control systems can be viewed as interconnections of spatially distributed multiple subsystems, where the individual subsystems share their information with each other through an underlying network that inherently introduces limitations on information flow. Inherent limitations on the flow of information among individual subsystems may stem from *structural constraints* of the network and/or *communication constraints* of the network. Hence, in order to design optimal control and estimation mechanisms for modern control systems, we must answer the following two practical but important questions:

(1) What are the fundamental communication limits to achieve a desired control performance and stability?

(2) What are the approaches one has to adopt to design a decentralized controller for a complex system to deal with structural constraints?

In this thesis, we consider four different problems within a game-theoretic framework to address the above questions.

The first part of the thesis considers problems of control and estimation with limited communication, which correspond to question (1) above.

We first consider the minimax estimation problem with intermittent observations. In this setting, the disturbance in the dynamical system as well as the sensor noise are controlled by adversaries, and the estimator receives the sensor measurements only sporadically, with availability governed by an independent and identically distributed (i.i.d.) Bernoulli process. This problem is cast in the thesis within the framework of stochastic zero-sum dynamic games. First, a corresponding stochastic minimax state estimator (SMSE) is obtained, along with an associated generalized stochastic Riccati equation (GSRE). Then, the asymptotic behavior of the estimation error in terms of the GSRE is analyzed. We obtain threshold-type conditions on the rate of

intermittent observations and the disturbance attenuation parameter, above which 1) the expected value of the GSRE is bounded from below and above by deterministic quantities, and 2) the norm of the sequence generated by the GSRE converges weakly to a unique stationary distribution.

We then study the minimax control problem over unreliable communication channels. The transmission of packets from the plant output sensors to the controller, and from the controller to the plant, are over sporadically failing channels governed by two independent i.i.d. Bernoulli processes. Two different scenarios for unreliable communication channels are considered. The first one is when the communication channel provides perfect acknowledgments of successful transmissions of control packets through a clean reverse channel, which is the TCP (Transmission Control Protocol), and the second one is when there is no acknowledgment, which is the UDP (User Datagram Protocol). Under both scenarios, the thesis obtains output feedback minimax controllers; it also identifies a set of explicit existence conditions in terms of the disturbance attenuation parameter and the communication channel loss rates, above which the corresponding minimax controller achieves the desired performance and stability.

In the second part of the thesis, we consider two different large-scale optimization problems via mean field game theory, which address structural constraints in the complex system stated in question (2) above.

We first consider two classes of mean field games. The first problem (P1) is one where each agent minimizes an exponentiated performance index, capturing risk-sensitive behavior, whereas in the second problem (P2) each agent minimizes a worst-case risk-neutral performance index, where a fictitious agent or an adversary enters each agent's state system. For both problems, a mean field system for the corresponding problem is constructed to arrive at a best estimate of the actual mean field behavior in various senses in the large population regime. In the finite population regime, we show that there exist ϵ-Nash equilibria for both P1 and P2, where the corresponding individual Nash strategies are decentralized as functions of the local state information. In both cases, the positive parameter ϵ can be taken to be arbitrarily small as the population size grows. Finally, we show that the Nash equilibria for P1 and P2 both feature robustness due to the risk-sensitive and worst-case behaviors of the agents.

In the last main chapter of the thesis, we study mean field Stackelberg differential games. There is one leader and a large number, say N, of followers. The leader holds a dominating position in the game, where he first chooses and then announces his optimal strategy, to which the N followers respond by playing a Nash game. The followers are coupled with each other through the mean field term, and are strongly influenced by the leader's strategy. From the leader's perspective, he is coupled with the N followers through the mean field term. In this setting, we characterize an approximated stochastic mean field process of the followers governed by the leader's strategy, which leads to a decentralized ϵ-Nash-Stackelberg equilibrium. As a consequence of decentralization, we subsequently show that the positive parameter ϵ can be picked arbitrarily small when the number of followers is arbitrarily large.

In the thesis, we also include several numerical computations and simulations, which illustrate the theoretical results.

To my wife and daughter, for their love and support

ACKNOWLEDGMENTS

First and foremost, I would like to express my deep gratitude to my adviser, Professor Tamer Başar, for his endless support and patience during my graduate studies. Over the years, he provided me with the freedom of my research, which allowed me to explore various interesting problems of control and games. His guidance and advice have helped me to reshape my research skills completely. It has been my great honor to have the opportunity to work with him for my Ph.D. research.

I would like to thank my doctoral committee members, Professor Bruce Hajek, Professor Geir Dullerud, and Professor Ali Belabbas. Their incisive and valuable comments during the preliminary exam have helped me to improve the exposition of the dissertation. The courses Random Processes and Game Theory by Professor Hajek, Robust Control and Control System Theory and Design by Professor Dullerud, and Stochastic Control by Professor Belabbas especially had a great impact on the contents of the thesis. I am also thankful to Professor Pierre Moulin for his courses Detection and Estimation Theory and Computational Inference and Learning, and Professor Renming Song for his course Probability Theory II: Stochastic Calculus and Differential Equations.

Furthermore, I want to express my sincere appreciation to my former adviser, Professor Chung Choo Chung. He was the first to introduce me to control theory, and has been a great mentor throughout my life. I learned a sincere research attitude from him. He also provided me with the opportunity to study for the Ph.D. degree, and helped me to receive the Fulbright scholarship for my Ph.D. studies.

I would also like to thank my friends, Dr. Won Hee Kim, Dr. Soo Bae Kim, Han Min Lee, Seok Jong Ryu, Rasoul Etesami, Xiaobin Gao, Abhishek Gupta, Ali Khanafer, and Khaled Alshehri for discussions about the thesis as well as our future life.

My graduate research was supported by the Fulbright Commission and the U.S. Air Force Office of Scientific Research MURI grant FA9550-10-1-0573. These funding sources are very much appreciated.

I could not have completed my Ph.D. journey without the support of my family. I would like to thank my parents, parents-in-law, and sister-in-law for their love and support.

Lastly, this thesis is dedicated to my wife and daughter, Mi Sung Park and Ji Yoo Moon. Their love and encouragement have helped me to finish my Ph.D. journey. I especially want to express my deep gratitude to my wife for her sacrifices and love.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SMSE	Stochastic Minimax State Estimator
GSRE	Generalized Stochastic Riccati Equation
MGRE	Modified Generalized Riccati Equation
ARE	Algebraic Riccati Equation
GARE	Generalized Algebraic Riccati Equation
RDE	Riccati Differential Equation
SDE	Stochastic Differential Equation
FBSDE	Forward-Backward Stochastic Differential Equation
i.i.d.	independent and identically distributed

NOTATIONS

$\mathbb{R}^n$	The set of n-dimensional real-valued vectors
$\mathbb{R}^{m\times n}$	The set of $m \times n$-dimensional real-valued matrices
$\mathbb{S}^n_{>0}$	The set of $n \times n$ positive definite matrices
$\mathbb{S}^n_{\geq 0}$	The set of $n \times n$ positive semi-definite matrices
x^T	Transpose of a vector x
X^T	Transpose of a matrix X
$\mathrm{Tr}(X)$	Trace of a matrix X
$\lvert x\rvert^2_S$	$x^T S x$ with $S \geq 0$ and a vector x (Chapters 2 and 3)
$\\|x\\|^2_S$	$x^T S x$ with $S \geq 0$ and a vector x (Chapters 4 and 5)
$\\|X\\|$	The induced 2-norm of a matrix X
$\\|\cdot\\|_\infty$	The supremum norm
$\rho(X)$	The spectral radius of a square matrix X
$\mathcal{C}^b_n$	The normed vector space of n-dimensional continuous and bounded functions with $\\|\cdot\\|_\infty$
ℓ^n_2	The space of square-summable sequences, taking values in $\mathbb{R}^n$
I_n	The $n \times n$ identity matrix
$\mathbf{1}_n$	The n-dimensional column vector whose elements are all 1
$\mathbf{0}_{n\times m}$	The $n \times m$-dimensional zero matrix
$\mathbf{0}_n$	$\mathbf{0}_n := \mathbf{0}_{n\times 1}$
$\mathbb{1}_{\{\cdot\}}$	The indicator function
$\mathrm{diag}\{\cdot\}$	A diagonal matrix with the argument

$\{x_k\}$	A sequence of vectors x_k, $k = 0, 1, ...,$ with appropriate dimensions
$\{X_k\}$	A sequence of matrices X_k, $k = 0, 1, ...,$ with appropriate dimensions
$\otimes$	Kronecker product
u_{-i}	$u_{-i} := \{u_1, ..., u_{i-1}, u_{i+1}, ..., u_N\}$
$\mathbb{E}$	Expectation operator
$\mathbb{P}$	Probability measure

CHAPTER 1

INTRODUCTION

The thesis focuses on various problems of control and estimation with limited information within a game-theoretic framework. The first part, that is Chapters 2 and 3, studies optimal estimation and control over unreliable communication channels within the stochastic zero-sum dynamic game framework. The second part, that is Chapters 4 and 5, considers two different large-scale optimization problems via mean field game theory.

In this chapter, we first introduce a general motivation on control and estimation with limited information, and provide an extensive literature review on the topic of the thesis. We then provide motivation for the questions studied in the thesis, and include a summary of the main results developed in the thesis.

1.1 Systems with Limited Information

In classical control systems, one considers designing a *centralized* controller that has access to the *entire* information on the system. In order to design a centralized controller, it is very common to assume that 1) the controller has access to information generated by the plant, and 2) the communication between the controller and the plant is flawless. With these idealized assumptions, a centralized controller can be designed using various optimal design methodologies to ensure stability and performance.

These idealized scenarios, however, cannot capture a large class of general real-world systems that are generally restricted to availability of *limited information* due to communication and/or structural constraints, which therefore reduces the ability of the control system to adapt to various uncertainties in the environment that affect its performance. For example, in remote control systems, the controller may not be able to communicate perfectly with the

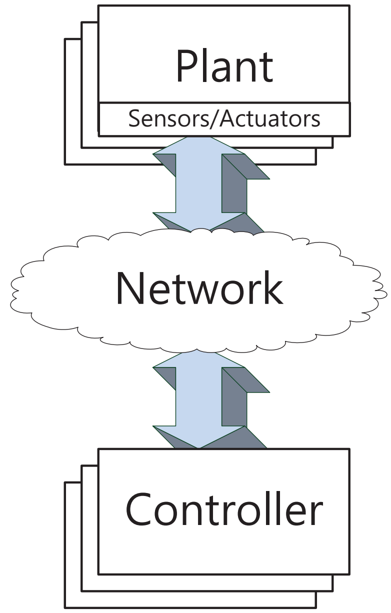


Figure 1.1: General networked control system configuration.

plant due to the unreliable nature of communication channels. Moreover, a centralized controller may not be implementable in large-scale systems, since the controller may not be able to access all the necessary information regarding the system.

In the following two subsections, we discuss in more detail the nature of limited information in control systems due to communication constraints and structural constraints, and these two constraints separately.

1.1.1 Systems with Communication Constraints

We first discuss communication constraints in control systems. What we mean by communication constraints is limitations on communication between the plant and the controller, and in both directions. An overall objective here is to study and characterize the degradation in performance in the face of such constraints, and to see how stability is affected. What are for example the fundamental limits on communication, in the same spirit as Shannon's information theory, beyond which the controller can stabilize the system, and performance degradation is within an acceptable range?

Over the past decade, the topic of *Networked Control Systems* (NCSs) has emerged to analyze the effect of communication limitations on control systems, which has led to a combined treatment of control and communications. A general configuration of NCSs is depicted in Fig. 1.1. In NCSs, the network characterizes communication limitations between controllers, plants, and sensors, and thus plays a crucial role in the achievement of the desired overall design objectives. Consequently, the main goal of research for NCSs is to analyze the behavior of a system under limited communication, and how to design controllers to cope with that limitation [1, 2, 3].

The application areas of NCSs are very broad; some of these applications are mobile sensor networks, remote surgery, unmanned aerial vehicles, and power and chemical plants [1, 4]. Command, control, communications, computers, and intelligence (C4I) systems in defense can also be seen as a sub-class of NCSs, where the local operation areas share their gathered information, and receive tactical operation messages from the headquarters through a wireless communication network that generally operates under some restrictions.

To capture the effects of communication limitations on NCSs, limitations such as delays, bounds on data rate, and packet drops are imposed on the communication network. These are natural because 1) delays are continually generated in communication while encoding, decoding, transmitting, and receiving signals, 2) every communication channel has a finite capacity within which the data has to be handled reliably, and 3) communication is unreliable, that is, it is subject to losses or link failures, especially in wireless transmissions [1, 2].

To better motivate this class of problems, consider a simple target estimation problem in a radar system, where estimation is performed based on sensor signals that are reflected from the target, as depicted in Fig. 1.2 [5, 6]. Due to a large volume of clutter and unforeseeable weather conditions, reflected signals can be lost or delayed randomly. In this case, after a certain number of consecutive measurement losses, the numbers of false alarms and/or missing targets increase, which results in losing the target information. Now, confronted with this scenario, the main objective should be to achieve stability and performance of the estimator in spite of measurement losses. To be specific, as mentioned earlier, we need to characterize the fundamental communication limits, above which the estimator would be able to

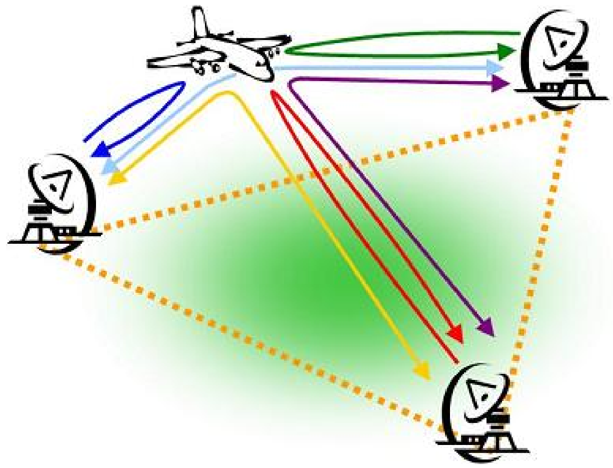


Figure 1.2: Configuration of the radar tracking system. Image courtesy: http://en.wikipedia.org

track the target reliably.

In Chapters 2 and 3 of the thesis, we focus on estimation and control problems over unreliable communication channels within the stochastic zero-sum dynamic game framework. A related literature review on this class of problems is provided in Section 1.2.1 of this chapter. Moreover, our motivation for Chapters 2 and 3, and a summary of the results, are given in Sections 1.3 and 1.4, respectively.

1.1.2 Systems with Structural Constraints

We now move on to structural constraints in control systems. A system that consists of several interconnected subsystems which are spatially distributed can be viewed as a large-scale system. In such systems, due to structural (or topological) constraints, it is not possible to physically connect the whole system; hence, the individual controllers will have access to limited information generated only from their close neighborhood [1, 3, 7]. Therefore, the information constraint in this case is generally imposed by the system structure, and a centralized approach with a single computer process to design an optimal controller is often not feasible. Moreover, even if the problem is mathematically tractable, the implementation of the controller suffers from computational complexity due to the high-dimensional state space with a large number of interconnected inputs and outputs [8, 9].

For example, in a communication network, there can be a large number of

Figure 1.3: The power network of USA. Image courtesy: http://www.foxnews.com

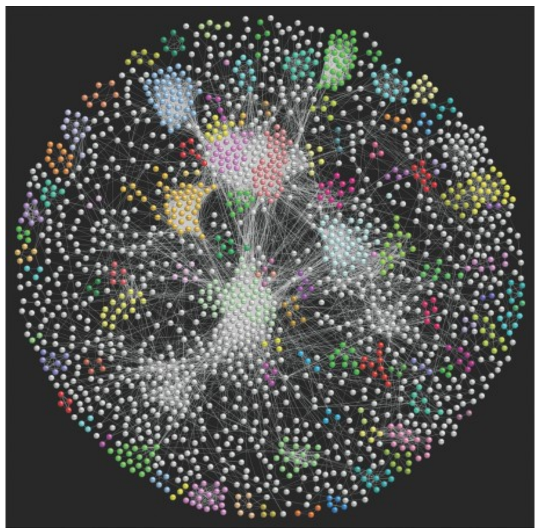


Figure 1.4: A large-scale map of protein interactions in fruit flies. Image courtesy: http://harvardmagazine.com

users who use limited resources provided by the service provider. Moreover, in the smart grid system, depicted in Fig 1.3, there are several utility companies and users, in which case for a large number of users, the traditional centralized optimization techniques suffer from computational complexity in attempting to manage the demand response [10]. Finally, the Web graph, various social networks with billions of vertices and trillions of edges, and biological networks as shown in Fig. 1.4, require computationally efficient algorithms to analyze their dynamics [11, 12].

To circumvent structural difficulties, and make the problem more tractable, a certain *localization* technique is needed to obtain a class of decentralized[1]

[1]Here, we do not distinguish between the definitions of "decentralized" and "distributed," since both terms reflect the localization in the large-scale system [8].

controllers so that the individual local controllers are able to generate an appropriate control action by using local measurements to control the large-scale system. But decentralized designs require much more involved analysis for the corresponding controller, since in some cases the decentralization can make the control problem intractable [13, 9].

In view of the above discussion, analyzing and optimizing large-scale systems constitute a real challenge. In Chapters 4 and 5, we study various large-scale optimization problems via mean field game theory, which is one way of coping with decentralization, localization, and complexity due to large scale. The problem description and literature review are provided in Section 1.2.2 of this chapter. Our motivation for Chapters 4 and 5 and a summary of the main results are provided in Sections 1.3 and 1.4, respectively.

1.2 Literature Review

In this section, we provide a review of the literature related to the topics studied in this thesis. We first review some of the previous work in NCSs with unreliable communication channels. We then discuss previous work on large-scale optimization and decentralized design via mean field game theory.

1.2.1 Optimal Control and Estimation over Unreliable Communication Channels

An initial study of optimal control over unreliable communication channels can be traced back to the uncertainty threshold principle developed by Athans, Ku and Gershwin in [14], which can be regarded as one of the first results that connect control and communication from Shannon's point of view. This result has been generalized to many different forms; see [15, 16], and the references therein. These results, however, were based on the assumption of perfect state measurement and hence cannot accommodate general communication limitations on a system.

Another way of looking at this class of problems is the context of Markov jump linear systems (MJLSs) [17, 18]. This modeling framework is appropriate, since many unreliable communication channels can be modeled simply by a two-state Markov chain with a transition probability distribution (or

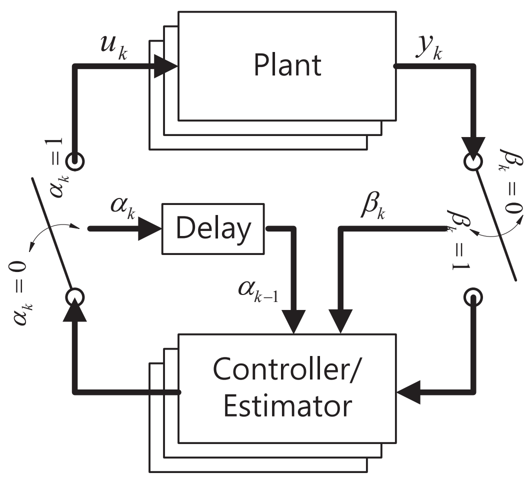


Figure 1.5: Control system configuration with TCP-networks.

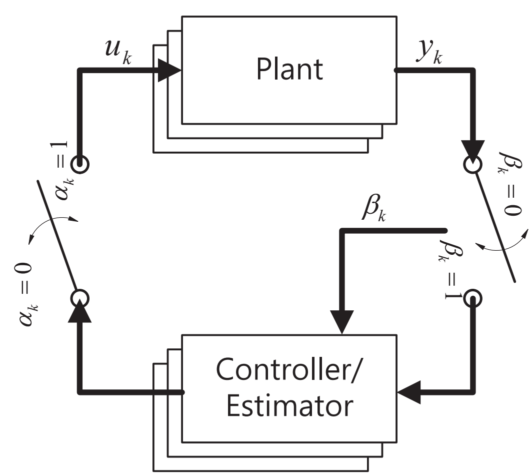


Figure 1.6: Control system configuration with UDP-networks.

rate). This approach, however, requires knowledge of the entire mode of the Markov chain (also known as the *form process*), which is not realistic for many practical control problems.

Subsequently, two different models of unreliable communication channels have been proposed in [19], where the unreliable communication channel was categorized based on whether the control packet reception is acknowledged (Transmission Control Protocol, TCP) or not (User Datagram Protocol, UDP); see Figs. 1.5 and 1.6, respectively. For this class of problems, the authors in [19] assumed that the quantities that model the unreliable communication channels, that is $\{\alpha_k\}$ and $\{\beta_k\}$ in Figs. 1.5 and 1.6, are independent and identically distributed (i.i.d.) Bernoulli processes.

In this setting, for the TCP-case, the authors in [19] considered the linear-quadratic Gaussian (LQG) problem, and showed that the separation principle holds, the LQG controller is linear in the measurement, and the stability region is determined by the unstable modes of the system and control and measurement loss rates, which can be viewed as the generalized uncertainty threshold principle. For the UDP-case, it was shown in [19] that the optimal controller is linear under some conditions; however, there is a *dual effect* between control and estimation [20].

The results in [19] were extended by [21] to the noisy measurement case. Specifically, for the TCP-case, it was shown that the LQG controller in [19] and the Kalman filter in [22] with the control input can be designed independently because there is no dual effect between filtering and control. The authors also characterized two independent critical values of control and measurement loss rates in terms of the unstable modes of the open-loop system. Furthermore, precise analytical bounds on these critical values were also provided. For the UDP-case, the authors showed that the optimal controller is generally nonlinear in the measurement, but is linear only for the perfect measurement case considered in [19]. The results for the TCP-case in [21] were extended to the multiple packet drop case in [23], to the generalized acknowledgment model in [24], and to the limited transmission bandwidth case in [25]. The decentralized LQG control problem over TCP-networks was discussed in [26], and the LQG problem with Markovian packet losses was studied in [27]. The stabilization issue and a characterization of the explicit critical values were also considered in [28, 29, 30].

As for H^∞ control over unreliable communication channels, the framework of Markov jump linear systems (MJLSs) was mostly used in the literature with some related references being [31, 32, 33, 34, 35, 36, 37]. Related to the MJLS approach, [38], [39], and [40] also studied H^∞ control with random packet losses by identifying a set of (different) linear matrix inequalities. The controllers, however, were restricted to be time invariant and are therefore suboptimal. It should be mentioned that with the framework of MJLSs, the information of the form process needs to be available to the controller, that is, the controller is mode dependent [36, 37]. Since the TCP-case does not

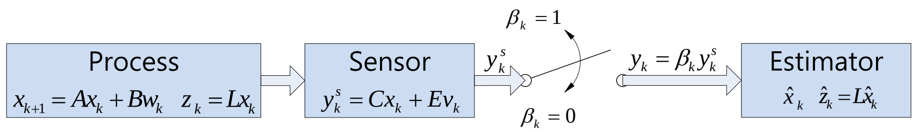


Figure 1.7: Configuration of estimation with intermittent observations.

have access to the current mode of the control packet loss information, as argued in [19], the unreliable channel modeled by MJLSs is not equivalent to the TCP-case (or for that matter to the UDP-case). One of the most recent sets of results on H^∞ control over unreliable communication channels was presented in [41]; however, the communication degradation there was due to packet delays rather than packet drops.

Along with control over unreliable communication channels, the estimation problem was studied extensively and independently. The general configuration of the underlying system is depicted in Fig. 1.7. This problem is also known as *estimation with intermittent observations*, which was initiated within the minimum mean-square estimation (MMSE), or Kalman filtering framework [42, 6].

Within the Kalman filtering framework, the first set of notable results were obtained by [22]. There, the stochastic Kalman filter and the associated stochastic Riccati equation (or the stochastic error covariance matrix), say P_k, were obtained such that their processes are dependent on the entire measurement arrival information, $\{\beta_k\}$. Moreover, [22] showed that there is a critical value of the measurement loss rate beyond which the expected value of the error covariance matrix, $\mathbb{E}\{P_k\}$, is bounded. It was also shown that this critical value is a function of the unstable modes of the system, and can be analyzed in terms of lower and upper bounds.

The difference between the Kalman filter in [22] and the Markov jump linear estimator (MJLE) in [43] is that the latter is dependent only on the current measurement arrival information β_k. It was shown in [22] that the Kalman filter in that paper is optimal over all possible estimators, and thus provides better estimation performance than the MJLE.

The results obtained in [22] were extended to many different forms, with some related references being [44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54]. In [44], a characterization of the critical value was studied when the system eigenvalues are distinct. In [45], it was shown that the lower bound of the critical value guarantees boundedness of $\mathbb{E}\{P_k\}$ if the system's observable space is invertible or the output matrix, say C, is invertible. For second-order systems, a closed-form expression for the critical value was obtained in [46]. In [48], Kalman filtering with two independent partial observation loss models was studied. Extended Kalman filtering (EKF) with intermittent observations, and its error covariance, were studied in [49], again in the expectation sense.

Instead of $\mathbb{E}\{P_k\}$, other performance metrics, or a correlated intermittent observation case, were studied in [47, 50, 51, 54, 53, 52] to provide other perspectives on the error covariance matrix, P_k. In [47], the authors considered boundedness of the upper envelope of P_k, and then showed its relationship to $\mathbb{E}\{P_k\}$ under a Markovian loss model. They also proved boundedness of $\mathbb{E}\{P_k^r\}$ for any $r \geq 1$ for scalar systems. In [50] and [51], the authors showed boundedness of P_k in the probabilistic sense, i.e., boundedness of $\mathbb{P}(P_k \leq M)$. In [53], more detailed results than in [47] were obtained when the intermittent observations are governed by a Markov process. Weak convergence of $\{P_k\}$ to a unique stationary distribution was discussed in [52] and [54]. While the author in [52] considered $\{P_k\}$ as a random walk and then proved weak convergence by using its mean contraction property, the authors in [54] showed the same result by modeling $\{P_k\}$ as an order-preserving and strongly sublinear random dynamical system.

The problem of H^∞ estimation with intermittent observations was studied in [55, 56, 57, 33, 58] within the framework of MJLSs. The estimators were restricted to be time invariant, since the MJLSs approach requires instantaneous information (that is β_k) on measurement losses. The authors in [59] proposed a time-varying H^∞ estimator; however, it is still suboptimal, since it uses the instantaneous measurement loss information.

1.2.2 Large-Scale Optimization via Mean Field Game Theory

Mean field games constitute a class of noncooperative stochastic differential (or dynamic) games, where there is a large number of players or agents, who interact with each other through a *mean field term* included in the individual performance indices and/or each agent's state system, which captures the average behavior of all agents. Rationality of each agent is a key underlying assumption in the characterization of the Nash equilibrium (or ϵ-Nash equilibrium) for the mean field game, which is the most general solution concept in noncooperative game theory [60, 61]. The mean field coupling term is also known as *mass behavior* or *macroscopic behavior* in statistical physics, since each individual agent can be viewed as a small particle; when there is a large number of particles, the effect of the individual particle on the mass behavior becomes negligible, and each particle follows this mass behavior via an individual-mass interaction consistency relationship.

Viewing each subsystem as an agent in a large-scale system, there are various applications of mean field games for multi-agent systems. In [62], the problem of a large number of coupled oscillators has been formulated within the mean field game framework, where decentralized optimal strategies were characterized to obtain an ϵ-Nash equilibrium. The problem of charging of a large number of plug-in electric vehicles was studied in [63] and [64]. In addition to this, there are several application domains of mean field games, such as engineering, finance, economics with a large number of firms, biology, etc.; see [65, 66, 67, 68, 69].

For games with a large number of agents, computing Nash equilibria for the corresponding game via direct methodologies as discussed in standard texts for dynamic games, such as [61], may be cumbersome and complicated, since complexity increases with the number and heterogeneity of the agents, and the dimension of the state space. Moreover, due to structural constraints imposed on the system, it is not realistic that each agent is able to access the state information generated by all other agents. It is more realistic that each agent is able to access only his own state information.

To resolve this difficulty, mean field analysis was used to obtain the best estimate of the actual mean field behavior, in terms of computationally low complexity optimal decentralized strategies that are functions of local information and constitute an ϵ-Nash equilibrium [70, 71, 72]. In [66], the authors developed independently a different approach to obtain the mean field equilibrium, applicable to more general models, which entails solving coupled forward-backward partial differential equations where the former is related to optimal control with the Hamilton-Jacobi-Bellman equation, and the latter is related to the mean field distribution with the Fokker-Planck-Kolmogorov equation. Both of these approaches are built on a platform that utilizes the fact that the impact of an individual agent on the mean field behavior becomes negligible when the number of agents goes to infinity.

Most of the mean field games discussed in the literature referenced above cannot capture robustness and risk averse-behavior, since the individual performance indices are risk-neutral. The class of risk-sensitive mean field games was introduced in [73] and [74], where the performance indices involve exponentiated integral cost. The one-agent version is equivalent to risk-sensitive optimal control, which was first introduced in [75] (see also [76, 77] and the references therein). In [75], the relationship of the linear-quadratic (LQ) risk-sensitive control problem to a LQ zero-sum differential game was also established. Later, the finite- and infinite-horizon risk-sensitive optimal control problems were considered in [78, 79, 80, 81, 82], and risk-sensitive differential nonzero-sum games were studied in [83].

In contrast to the class of mean field games referenced above, [84] and [85] considered the situation when there are one major agent and a large number of minor agents, in view of which stochastic mean field approximation was introduced, and ϵ-Nash equilibria were obtained. Specifically, due to strong influence of the major agent on minor agents, the approximated mean field coupling term is no longer deterministic, but a stochastic process driven by the Brownian motion of the leader. In [84], the state augmentation method was developed via the strong law of large numbers to characterize the best stochastic mean field process when the followers are heterogeneous with K distinct models. In [85], fixed point analysis was applied to obtain similar results as in [84] when the dynamics and costs of the followers are parametrized within a continuum set. These two different approaches lead to (different) decentralized optimal strategies for the individual agents that constitute (different) ϵ-Nash equilibrium. The nonlinear counterpart of mean field games with major and minor agents was studied in [86].

Finally, it should be mentioned that mean field games discussed above are *Nash* games. That is, each agent determines his optimal strategy noncooperatively and all simultaneously, which leads to ϵ-Nash equilibria, and there is no hierarchy of decision making between the agents. On the other hand, if one wants to model a certain hierarchical structure in mean field games, the corresponding problem can be formulated by employing the *Stackelberg* setting. Classical Stackelberg games are hierarchical decision-making problems, where there is a leader with a dominant position over the follower [87].

The leader first announces his optimum strategy by taking into account the rational reactions of the followers. The follower then chooses his optimal strategy based on the leader's strategy, and finally the leader comes back and implements his announced strategy, thus generating his action. When there is such a solution, the resulting optimum strategies for the leader and the follower form a Stackelberg equilibrium for the corresponding game [61].

Stackelberg differential and dynamic games have been studied extensively in the literature since 1970, and detailed expositions can be found in [61, 88, 89, 90, 91, 92, 93], and the references therein. Stackelberg games have a wide range of applications. For example, in the setting of communication networks, one can have a single service provider and a (large) number of users, where the service provider sets the usage price(s) for the Nash followers [94]. In the smart grid, optimal demand response management can be studied within the framework of Stackelberg games, where the utility companies are leaders, and the users are followers [95].

1.3 Motivation for the Thesis

Chapters 2 and 3 study NCSs with unreliable communication channels. Our motivation for Chapters 2 and 3 is as follows:

- The problem of optimal estimation with intermittent observations captured in Fig. 1.7 was investigated exclusively within the Kalman filtering framework; however, the worst-case scenario of this class of problems (which captures robustness) had not yet been addressed until we initiated work on this topic in [96, 97]. Some available results on H^∞ estimation with intermittent observations are suboptimal due to their restrictive information and estimator structure.

- The available results on optimal control over TCP- and UDP-like communication channels captured in Figs. 1.5 and 1.6 had been obtained within the LQG framework, and the same problem for robust control within the worst-case scenario had not yet been addressed until our work initiated it in [98, 99, 100, 101]. Some available results on H^∞ control over unreliable communication channels discussed are suboptimal, and they are not related to the TCP- and the UDP-cases due to their specified information structures.

In Chapters 4 and 5, we consider large-scale optimization problems under structural constraints via mean field game theory. Our motivation for Chapters 4 and 5 is as follows:

- Following up with [73] and [74], it is a natural (but challenging) next step to extend the previous work on risk-sensitive mean field games to the heterogeneous agent case with infinite-horizon cost functions. Moreover, it is necessary to study the robustness property of the corresponding ϵ-Nash equilibria in view of the fact that the risk-sensitive approach is closely connected with H^∞ control. Finally, we also need to study the relationship between risk-sensitive, worst-case, and risk-neutral mean field games with respect to some design parameters, as was done for classical one-agent optimal control problems in earlier literature, for example [75, 78, 77, 79].

- As discussed in Section 1.2.2, the previous work on mean field games are Nash games. Hence, it is necessary to study hierarchical decision-making problems via mean field game theory within the Stackelberg framework.

1.4 Main Results of the Thesis

We state a summary of the main results developed in this thesis:

- In Chapter 2, within the setting of intermittent observations depicted in Fig. 1.7, we obtain a stochastic minimax state estimator (SME) with an associated generalized stochastic Riccati equation (GSRE) by solving a corresponding stochastic zero-sum dynamic game. We then study the asymptotic behavior of the estimation error in terms of the GSRE, and show that under some conditions, 1) the GSRE is bounded below and above in the expectation sense, and 2) the GSRE converges weakly to a unique stationary distribution. This chapter is based on [96, 97].

- In Chapter 3, we obtain classes of output feedback minimax controllers for both the TCP- and the UDP-cases by solving a corresponding stochastic zero-sum dynamic game. We identify a set of explicit threshold conditions in terms of the loss rates and the disturbance attenuation parameter, above which the corresponding minimax controller exists, and achieves the desired disturbance attenuation performance and stability. This chapter is based on [98, 99, 100, 101].

- In Chapter 4, we consider risk-sensitive and robust mean field games. For both problems, we obtain ϵ-Nash equilibria, where the individual Nash strategies are decentralized as a function of local state information, and the positive parameter ϵ can be picked arbitrarily close to zero when the number of agents is arbitrarily large. We show that the ϵ-Nash equilibria for both problems are *partially* equivalent. Finally, we also show that the ϵ-Nash equilibria for both problems feature robustness due to the risk-sensitive and worst-case behaviors of the agents. This chapter is based on [102, 103].

- In Chapter 5, we study mean field Stackelberg differential games with one leader and a large number of followers. Given an arbitrary strategy of the leader, we characterize an approximated stochastic mean field process, and obtain an ϵ-Nash equilibrium for the followers, where the individual Nash strategies are decentralized as a function of local state information. We then obtain a decentralized (ϵ_1, ϵ_2)-Stackelberg strategy for the leader, where we show that the positive parameters ϵ_1 and ϵ_2 converge to zero when the population size of the followers grows to infinity. This chapter is based on [104, 105].

1.5 Organization of the Thesis

The outline of the thesis is as follows. In Chapter 2, we study the minimax estimation problem with intermittent observations. The minimax control problem for both the TCP- and UDP-cases is presented in Chapter 3. In Chapter 4, we consider risk-sensitive and robust mean field games. We study mean field Stackelberg differential games in Chapter 5. We collect our concluding remarks in Chapter 6, where some future research directions are also provided. The appendices contain proofs and important lemmas that are used to prove the mains results presented in the thesis.

Part I

Control and Estimation over Unreliable Communication Channels

CHAPTER 2

MINIMAX ESTIMATION WITH INTERMITTENT OBSERVATIONS

2.1 Introduction

We study the minimax estimation problem with intermittent observations, as depicted in Fig. 1.7. As mentioned in Chapter 1, although there has been significant progress on the intermittent observation problem within the Kalman filtering framework, it has not yet been addressed thoroughly through the worst-case or H^{∞} approach. Some relevant results for the problem of H^{∞} estimation with intermittent observations were obtained in [55, 56, 57, 33], and [58], where different sets of linear matrix inequalities were derived for the H^{∞} performance. These results, however, are related more to the theory of Markov jump linear estimators (MJLEs) and are therefore suboptimal, since the estimators are restricted to be time-invariant and obtained under the assumption of instantaneous measurement arrival.

In this chapter, by formulating the problem within the framework of stochastic zero-sum dynamic games, we first obtain a stochastic minimax state estimator (SMSE) and an associated generalized stochastic Riccati equation (GSRE), both of which are time varying and random, and are dependent on the sequence of the random measurement arrival information $\{\beta_k\}$ and the H^{∞} disturbance attenuation parameter γ. We then identify an existence condition for the SMSE in terms of the GSRE and γ. We also show that under that existence condition, the SMSE is able to attenuate arbitrary disturbances within the level of γ. Moreover, we show that for the extreme scenario that corresponds to least disturbance attenuation (that is, as $\gamma \to \infty$), the SMSE and the GSRE converge, respectively, to the Kalman filter and its stochastic Riccati equation $\{P_k\}$ in [22].

The second objective of this chapter is to analyze the asymptotic behavior of the estimation error in terms of the GSRE. In particular, we prove boundedness of the sequence generated by the GSRE in the expectation sense, and also show its weak convergence. More specifically, we first show that under the existence condition, there exist both a critical value of the measurement loss rate and a critical value for the disturbance attenuation parameter beyond which the expected value of the sequence generated by the GSRE can be bounded both below and above. Second, we prove that under the existence condition, the norm of the sequence generated by the GSRE converges weakly to a unique stationary distribution. For both cases, we show that when $\gamma \to \infty$, the corresponding asymptotic results are equivalent to those in [22] and [54]. We also demonstrate by simulations that the SMSE outperforms the stationary and suboptimal H^∞ MJLE in [58].

Organization

The structure of the chapter is as follows. In Section 2.2, we formulate the problem of minimax estimation with intermittent observations. In Section 2.3, we obtain the SMSE and GSRE, and characterize the existence condition. In Section 2.4, we analyze the asymptotic behavior of the GSRE. In Section 2.5, we present simulation results. We end the chapter with the concluding remarks of Section 2.6.

2.2 Problem Formulation

Consider the following linear dynamical system:

$$x_{k+1} = Ax_k + Dw_k \tag{2.1a}$$

$$y_k^s = Cx_k + Ev_k \tag{2.1b}$$

$$y_k = \beta_k y_k^s \tag{2.1c}$$

$$z_k = Lx_k, \tag{2.1d}$$

where $x_k \in \mathbb{R}^n$ is the state; $\{w_k\} \in \ell_2^p$ and $\{v_k\} \in \ell_2^m$ are the disturbance input and the measurement noise sequences, respectively; $y_k^s \in \mathbb{R}^m$ is the sensor output, y_k is the channel output that is available to the estimator (see Fig. 1.7); $z_k \in \mathbb{R}^q$ is the variable that needs to be estimated; and A, C, D,

E, L are time-invariant matrices with appropriate dimensions. We assume that (A, C) is observable and (A, D) is controllable. We also assume that E is square and non-singular, and define $V := EE^T$. In (2.1), the sequence of random variables, $\{\beta_k\}$, models intermittency of observations between the sensor and the estimator (see Fig. 1.7), which is an i.i.d. Bernoulli process with $\mathbb{P}(\beta_k = 1) = \beta$.

Define the information that is available to the estimator at each time k as:

$$\mathcal{I}_k := \{\beta_{0:k}, y_{0:k}\}. \tag{2.2}$$

It is worth noting that the H^∞ MJLE in [55] and [58] utilizes the partial information, $\{y_{0:k}, \beta_k\}$, which is a subset of the information structure in (2.2).

Having the information structure as in (2.2), we seek an estimate $\hat{x}_k$ (or $\hat{z}_k$) of the actual state x_k, to be generated by $\hat{x}_k = \pi_k(\mathcal{I}_k)$, where π_k is an admissible (Borel measurable) estimator policy to be determined. We denote the class of admissible estimation policies by Π.

Now, our first main objective in this chapter is to find a recursive estimator policy $\pi \in \Pi$ under (2.2) that minimizes the following worst-case cost function:

$$\ll \mathcal{T}_\pi^N \gg := \sup_{x_0, w_{0:N-1}, v_{0:N-1}} \frac{\mathbb{E}\{\sum_{k=0}^{N-1} |z_k - \hat{z}_k|_Q^2\}^{1/2}}{\mathbb{E}\{|x_0 - \tilde{x}_0|_{Q_0}^2 + \sum_{k=0}^{N-1} |w_k|^2 + \beta_k |v_k|^2\}^{1/2}}, \tag{2.3}$$

where $Q \in \mathbb{S}^n_{\geq 0}$, $Q_0 \in \mathbb{S}^n_{>0}$, and $\tilde{x}_0 \in \mathbb{R}^n$ is a known bias term which stands for some initial estimate of x_0. Note that in (2.3), the expectation is taken with respect to $\{\beta_k\}$, since the estimator policy that we are seeking should be dependent on $\{\beta_k\}$ in view of our information structure given in (2.2). The problem in (2.3) can be regarded as an H^∞ estimation problem [106, 107].

By invoking the formulation of the corresponding soft-constrained game, the cost function of the associated zero-sum dynamic game parametrized by the disturbance attenuation parameter, $\gamma > 0$, can be written as follows (see [61]):

$$\begin{aligned} &J_\gamma^N(\pi, x_0, w_{0:N-1}) \qquad (2.4) \\ &= \mathbb{E}\Big\{-\gamma^2 |x_0 - \tilde{x}_0|_{Q_0}^2 + \sum_{k=0}^{N-1} |x_k - \hat{x}_k|_Q^2 - \gamma^2\big(|w_k|^2 + |y_k - \beta_k C x_k|_{V^{-1}}^2\big)\Big\}, \end{aligned}$$

where (2.1c) is used with $v_k = E^{-1}(y_k - \beta_k C x_k)$, and $Q := L^T Q L \geq 0$.[1]

In view of the deterministic H^∞ estimation problem considered in [106], if (2.4) admits a saddle-point solution[2] for some γ, then the corresponding estimator policy, $\pi_\gamma^* \in \Pi$, is a minimax policy that is parametrized by γ, and under this policy, we have the bound $\ll \mathcal{T}_{\pi_\gamma^*}^N \gg \leq \gamma$. Moreover, when γ is sufficiently large or γ goes to infinity, the minimax policy (or estimator) exists, which corresponds to the Kalman filter [106].

Bearing in mind the previous observation, we seek a minimax estimator policy, π_γ^*, that solves

$$\inf_{\pi \in \Pi} \sup_{x_0, w_{0:N-1}} J_\gamma^N(\pi, x_0, w_{0:N-1}) = \sup_{x_0, w_{0:N-1}} J_\gamma^N(\pi_\gamma^*, x_0, w_{0:N-1}). \tag{2.5}$$

Such a π_γ^* will be obtained in Section 2.3. We also need to characterize the smallest values of γ and β, say γ_s^* and β_s, for which the minimax estimator exists for all $\gamma > \gamma_s^*$ and $\beta > \beta_s$. Obviously, γ_s^* and β_s are coupled with each other; hence they cannot be determined independently. Moreover, γ_s^*, provided that it is finite, will be the value of (2.3).

After solving (2.5) and characterizing γ_s^* and β_s, or more precisely, finding the relationship between γ_s^* and β_s, we will analyze the asymptotic behaviors of the minimax estimator in Section 2.4.

2.3 Stochastic Minimax State Estimator

In this section, we obtain a minimax estimator by solving the zero-sum dynamic game in (2.4). The main result is the following.

Theorem 2.1. *Consider the zero-sum dynamic game (2.4) subject to (2.1) with $k \in [0,\ N-1]$, $\beta \in [0,1]$, and a fixed $\gamma > 0$. Then:*

[1] We can also use the measurement model $y_k = \beta_k C x_k + E v_k$ with

$$\ll \mathcal{T}_\pi^N \gg := \sup_{x_0, w_{0:N-1}, v_{0:N-1}} \frac{\mathbb{E}\{\sum_{k=0}^{N-1} |z_k - \hat{z}_k|_Q^2\}^{1/2}}{\mathbb{E}\{|x_0 - \tilde{x}_0|_{Q_0}^2 + \sum_{k=0}^{N-1} |w_k|^2 + |v_k|^2\}^{1/2}},$$

to arrive at the same zero-sum dynamic game as in (2.4). Note that the above model considers the situation when the estimator receives *noise only signal* when there are measurement losses.

[2] See [106] and [61] for the definition of saddle point for zero-sum dynamic games.

(i) There exists a stochastic minimax state estimator (SMSE) if

$$\rho(\Sigma_k Q) < \gamma^2, \text{ almost surely (a.s.) for all } k, \tag{2.6}$$

where Σ_k with $\Sigma_0 = Q_0^{-1}$ is generated by a generalized stochastic Riccati equation (GSRE)

$$\Sigma_{k+1} = A(\Sigma_k^{-1} - \gamma^{-2}Q + \beta_k C^T V^{-1} C)^{-1} A^T + DD^T. \tag{2.7}$$

(ii) The SMSE is given by

$$\begin{aligned} \hat{z}_k &= L\hat{x}_k \\ \hat{x}_{k+1} &= A\hat{x}_k + \beta_k A K_k (y_k - C\hat{x}_k), \ \hat{x}_0 = \tilde{x}_0, \end{aligned} \tag{2.8}$$

where K_k is the estimator gain that can be written as

$$K_k = (\Sigma_k^{-1} - \gamma^{-2}Q + \beta_k C^T V^{-1} C)^{-1} C^T V^{-1}. \tag{2.9}$$

(iii) Suppose that (2.6) holds for all k. Then the SMSE achieves the performance level of γ, that is, $\ll \mathcal{T}^N_{\pi^*_\gamma} \gg \leq \gamma$.

Proof. Since we seek a causal estimator, by using forward dynamic programming, we can introduce the quadratic *cost-to-come (worst past cost)* function $W_k(x_k) = \mathbb{E}\{-|x_k - \bar{x}_k|^2_{Z_k} + l_k | \mathcal{I}_k\}$, where $Z_k > 0$, $Z_0 = \gamma^2 Q_0$, and $l_0 = 0$ [106, 107, 96]. Then, the cost from the initial state to stage $k+1$ is

$$\begin{aligned} &\mathbb{E}\{|x_{k+1} - \bar{x}_{k+1}|^2_{Z_{k+1}} - l_{k+1} | \mathcal{I}_{k+1}\} \\ &= \min_{(w_k, x_k)} \left[-|x_k - \hat{x}_k|^2_Q + \gamma^2 |w_k|^2 + \gamma^2 |y_k - \beta_k C x_k|^2_{V^{-1}} + |x_k - \bar{x}_k|^2_{Z_k} - l_k \right], \end{aligned}$$

where the equality follows from the definition of the information structure $\mathcal{I}_{k+1}$.

For existence of a unique minimizer, by Lemma 6.1 in [106], given $\mathcal{I}_k$, it is necessary to have $Z_k - Q > 0$ for all k. Then, by Lemma 6.2 in [106], the minimum cost at stage $k+1$ can be obtained as

$$\begin{aligned}
&\mathbb{E}\{|x_{k+1} - \bar{x}_{k+1}|^2_{Z_{k+1}} - l_{k+1}|\mathcal{I}_{k+1}\} \\
&= |x_{k+1} - AU_k^{-1}(Z_k\bar{x}_k + \beta_k\gamma^2 C^T V^{-1} y_k - Q\hat{x}_k)|^2_{\bar{Z}_k^{-1}} \\
&\quad + |Q\hat{x}_k - Z_k\bar{x}_k - \beta_k\gamma^2 C^T V^{-1} y_k|^2_{U_k^{-1}} + |\bar{x}_k|^2_{Z_k} - |\hat{x}_k|^2_Q + \gamma^2|y_k|^2_{V^{-1}} - l_k,
\end{aligned}$$

where $U_k := Z_k + \beta_k\gamma^2 C^T V^{-1} C - Q$ and $\bar{Z}_k := AU_k^{-1}A^T + \gamma^{-2}DD^T$. Note that the last equation is equivalent to its conditional expectation given $\mathcal{I}_{k+1}$. Since this is true for all k, the dynamic equation for $\bar{x}_k$ can be written as

$$\begin{aligned}
\bar{x}_{k+1} &= AU_k^{-1}(Z_k\bar{x}_k + \beta_k\gamma^2 C^T V^{-1} y_k - Q\hat{x}_k) \\
Z_{k+1} &= (AU_k^{-1}A^T + \gamma^{-2}DD^T)^{-1}.
\end{aligned} \tag{2.10}$$

Also, with $Z_0 = \gamma^2 Q_0$, let $\Sigma_k := \gamma^2 Z_k^{-1}$; then (2.10) can be rewritten as

$$\bar{x}_{k+1} = A\bar{x}_k + \beta_k AK_k((y_k - C\bar{x}_k) + \gamma^{-2}Q(\bar{x}_k - \hat{x}_k)).$$

Now, choose the estimator policy to be the certainty equivalence policy, i.e., $\hat{x}_k = \bar{x}_k$ for all k [106, 107], which can be obtained by solving the following optimization problem:

$$\hat{x}_k = \arg\max_{x_k} \mathbb{E}\{-\gamma^2|x_k - \bar{x}_k|^2_{\Sigma_k^{-1}} + l_k|\mathcal{I}_k\}.$$

Note that the above problem has a unique solution. Then we have (2.7) and (2.8).

To prove part (iii), observe that under (2.6), by the definition of the cost-to-come function [107], $W_k(x_k) \leq 0$ for all disturbances $(x_0, w_{0:N-1})$ and for all k. This implies that the value of the zero-sum game (2.4) is finite and bounded from above by zero. Hence, the SMSE achieves the disturbance attenuation level γ. Note that if (2.6) does not hold at $\bar{k} \in [0, N-1]$, the zero-sum game is unbounded because $W_k(x_k)$ can be made arbitrarily large by an appropriate choice of the disturbance [106, 107]. This implies that γ cannot be the disturbance attenuation level for the SMSE. This completes the proof of the theorem. □

We define

$$\Gamma(\beta) := \{\gamma > 0 \ : \ \rho(\Sigma_k Q) < \gamma^2, \text{ a.s. } \forall k \in [0,\infty)\}$$
$$\Lambda(\gamma) := \{\beta \in [0,1) \ : \ \rho(\Sigma_k Q) < \gamma^2, \text{ a.s. } \forall k \in [0,\infty)\}.$$

Let $\gamma_s^*(\beta) := \inf\{\gamma \ : \ \gamma \in \Gamma(\beta)\}$ and $\beta_s(\gamma) := \inf\{\beta \ : \ \beta \in \Lambda(\gamma)\}$. Then we have the following proposition which follows from the definitions of $\gamma_s^*(\beta)$ and $\beta_s(\gamma)$.

Proposition 2.1. *Suppose that γ is finite, and $\gamma > \gamma_s^*(\beta)$ and $\beta > \beta_s(\gamma)$. Then $\rho(\Sigma_k Q) < \gamma^2$ holds a.s. for all k.* □

We have two observations on $\gamma_s^*(\beta)$. First, when A is stable, $\gamma_s^*(\beta)$ is finite for all $\beta \in [0,1]$. This is because when $\beta = 0$, which is the worst-case communication channel, the problem reduces to the open-loop H^∞ estimation problem of the stable system, and it was shown in [106] that $\gamma_s^*(0)$ is finite. Second, when A is unstable, $\gamma_s^*(0)$ is not finite since $\Gamma(0)$ is empty. Notice that $\gamma_s^*(1)$ is finite since (2.7) becomes the Riccati equation of the deterministic H^∞ estimation problem in [106].

We have some remarks. Note that P_k below is the error covariance matrix of the Kalman filter in [22], which is provided in (A.14) (or (A.15)).

Remark 2.1. *(i) $\gamma_s^*(\beta)$ and $\beta_s(\gamma)$ are functions of each other and therefore cannot be determined independently.*

(ii) By induction, we can show that $P_k \leq \Sigma_k$ a.s. for all k, provided that the condition in Proposition 2.1 holds and $P_0 = Q_0^{-1}$.

(iii) As $\gamma \to \infty$, $\Sigma_k \to P_k$ and the SMSE converges to the Kalman filter in (A.13).

(iv) By using the matrix inversion lemma, the GSRE can be written as

$$\Sigma_{k+1} = A\Sigma_k A^T + DD^T - A\Sigma_k \begin{pmatrix} \beta_k C^T & G^T \end{pmatrix} \times \begin{pmatrix} V + \beta_k C\Sigma_k C^T & \beta_k C\Sigma_k G^T \\ \beta_k G\Sigma_k C^T & G\Sigma_k G^T - \gamma^2 I \end{pmatrix}^{-1} \begin{pmatrix} \beta_k C \\ G \end{pmatrix} \Sigma_k A^T,$$

where $G^T G = Q$. Then clearly, $\Sigma_k \to P_k$ as $\gamma \to \infty$.

Before concluding this section, it is worth noting that H^∞ MJLEs in [55] and [58] utilize the partial information structure for estimation; hence, they are time-invariant and suboptimal. In fact, the SMSE will provide better estimation performance than H^∞ MJLEs. This fact will be demonstrated by numerical examples in Section 2.5.2.

2.4 Asymptotic Analysis of the GSRE

In this section, we provide an asymptotic analysis of the GSRE. The first subsection deals with boundedness of $\mathbb{E}\{\Sigma_k\}$. This constitutes a generalization of Result A.1 in Appendix A.3. The second subsection considers the weak convergence of $\{\|\Sigma_k\|\}$. This result can be seen as a minimax counterpart of Theorem 3.1 in [54], which considers the weak convergence of $\{\|P_k\|\}$ where P_k is the SRE given in Appendix A.3 as (A.14) (or (A.15)).

2.4.1 Boundedness of $\mathbb{E}\{\Sigma_k\}$

Consider the following modified Lyapunov equation and modified generalized Riccati equation (MGRE):

$$\breve{\Sigma}_{k+1} = (1-\beta)h_1(\gamma, \breve{\Sigma}_k) \tag{2.11}$$

$$\bar{\Sigma}_{k+1} = (1-\beta)h_1(\gamma, \bar{\Sigma}_k) + \beta h_2(\gamma, \bar{\Sigma}_k) =: h(\gamma, \beta, \bar{\Sigma}_k), \tag{2.12}$$

where $\breve{\Sigma}_0 = \bar{\Sigma}_0 = Q_0^{-1}$, and the functions h_1 and h_2 are defined in Appendix A.1 as (A.2a) and (A.2b), respectively. The objective here is to show that under some conditions, (2.11) and (2.12) constitute respectively lower and upper bounds on $\mathbb{E}\{\Sigma_k\}$.

Proposition 2.2. *Introduce the following algebraic Riccati equation (ARE):*

$$\breve{\Sigma} = (1-\beta)h_1(\gamma, \breve{\Sigma}). \tag{2.13}$$

Define

$$\breve{\Gamma} := \{\gamma > 0 : \ \rho(\breve{\Sigma}^+ Q) < \gamma^2, \ \breve{\Sigma}^+ \in \mathbb{S}^n_{>0} \ \textit{solves (2.13)}\}.$$

Let $\breve{\gamma}^* := \inf\{\gamma : \ \gamma \in \breve{\Gamma}\}$ *and* $\breve{\beta}_c := 1 - \frac{1}{\rho^2(A)}$. *Then if* γ *is finite, and* $\gamma > \breve{\gamma}^*$ *and* $\beta > \breve{\beta}$, *we have* $\{\breve{\Sigma}_k\} \to \breve{\Sigma}^+$ *as* $k \to \infty$.

Proof. See Appendix A.1. □

Proposition 2.3. *Introduce the following modified generalized algebraic Riccati equation (MGARE):*

$$\bar{\Sigma} = h(\gamma, \beta, \bar{\Sigma}) = (1-\beta)h_1(\gamma, \bar{\Sigma}) + \beta h_2(\gamma, \bar{\Sigma}). \tag{2.14}$$

Define the following sets:

$$\begin{aligned}
\bar{\Gamma}(\beta) &:= \{\gamma > 0 : \ \rho(\bar{\Sigma}Q) < \gamma^2, \bar{\Sigma} \in \mathbb{S}^n_{>0} \ \textit{solves (2.14)}\} \\
\bar{\Lambda}(\gamma) &:= \{\beta \in [0,1) : \rho(\bar{\Sigma}Q) < \gamma^2, \bar{\Sigma} \in \mathbb{S}^n_{>0} \ \textit{solves (2.14)}\}.
\end{aligned}$$

Let $\bar{\gamma}^*(\beta) := \inf\{\gamma : \ \gamma \in \bar{\Gamma}(\beta)\}$ *and* $\bar{\beta}_c(\gamma) := \inf\{\beta : \ \beta \in \bar{\Lambda}(\gamma)\}$. *Suppose* $Q_0^{-1} \leq DD^T$. *Then, for any finite* $\gamma > \bar{\gamma}^*(\beta)$ *and* $\beta > \bar{\beta}_c(\gamma)$, *as* $k \to \infty$, $\{\bar{\Sigma}_k\} \to \bar{\Sigma}^+$ *where* $\bar{\Sigma}^+$ *is a fixed point of the MGARE with* $\rho(\bar{\Sigma}^+Q) < \gamma^2$.

Proof. See Appendix A.1. □

Proposition 2.4. *Suppose that the condition in Proposition 2.3 holds. Then:*

(i) Suppose $\bar{\Gamma}(\beta)$ *is not empty with* β_1 *and* β_2. *If* $\beta_1 \geq \beta_2$, *then* $\bar{\gamma}^*(\beta_1) \leq \bar{\gamma}^*(\beta_2)$. *Also* $\bar{\gamma}^*(\beta) \geq \bar{\gamma}^*(1)$ *for all* β.

(ii) If $\gamma_1 \geq \gamma_2 > \bar{\gamma}^*(\beta)$ *is finite, then* $\bar{\beta}_c(\gamma_1) \leq \bar{\beta}_c(\gamma_2)$. *Also, as* $\gamma \to \infty$, $\bar{\beta}_c(\gamma) \to \bar{\lambda}$ *where* $\bar{\lambda}$ *is defined in Result A.1 in Appendix A.3.*

Proof. See Appendix A.1. □

We now show the existence of critical values, $\beta_c(\gamma) \in [0,1)$ and $\gamma_c^*(\beta) > 0$, which determine boundedness of $\mathbb{E}\{\Sigma_k\}$.

Theorem 2.2. *Suppose that the condition in Proposition 2.1 holds. Then, there exist* $\beta_c(\gamma) \in [0,1)$ *and* $\gamma_c^*(\beta) > 0$ *such that*

$$\begin{aligned}
&\forall k, \ \mathbb{E}\{\Sigma_k\} \leq M(\gamma, \Sigma_0), && \textit{if } \beta > \beta_c(\gamma) \textit{ and } \gamma > \gamma_c^*(\beta) \\
&\lim_{k\to\infty} \mathbb{E}\{\Sigma_k\} = \infty, && \textit{otherwise,}
\end{aligned}$$

where $M(\gamma, \Sigma_0)$ *depends on the initial condition of the GSRE and* γ.

Proof. When A is stable, there exists $\gamma > 0$ such that $\mathbb{E}\{\Sigma_k\}$ is bounded for all $\beta \in [0,1]$. This is because the GSRE becomes the deterministic generalized Riccati equation of the H^∞ estimation problem when $\beta = 1$, and is the modified Lyapunov equation when $\beta = 0$ where the boundedness of the former and the latter were shown in [106] and Appendix A.1, respectively.

Consider the case when A is unstable. If $\beta = 1$, then under the controllability and observability assumptions, $\mathbb{E}\{\Sigma_k\}$ converges to a positive definite matrix. If $\beta = 0$, then the problem is equivalent to the open-loop estimation; therefore, $\mathbb{E}\{\Sigma_k\} \to \infty$ as $k \to \infty$. Now, suppose that there is β_1 such that $\mathbb{E}\{\Sigma_k\}$ is bounded. Clearly, by Lemma A.3(i) in Appendix A.2, $\mathbb{E}\{\Sigma_k\}$ is bounded for all $\beta > \beta_1$. In fact, if $\beta_2 \geq \beta_1$, then we have

$$\begin{aligned}\mathbb{E}_{\beta_1}\{\Sigma_{k+1}\} &= (1-\beta_1)\mathbb{E}\{h_1(\gamma,\Sigma_k)\} + \beta_1\mathbb{E}\{h_2(\gamma,\Sigma_k)\} \\ &\geq (1-\beta_2)\mathbb{E}\{h_1(\gamma,\Sigma_k)\} + \beta_2\mathbb{E}\{h_2(\gamma,\Sigma_k)\} = \mathbb{E}_{\beta_2}\{\Sigma_{k+1}\},\end{aligned}$$

where Lemma A.3(i) is used to arrive at the inequality. It is possible to define $\beta_c(\gamma)$ as follows:

$$\beta_c(\gamma) := \inf\{\beta \in [0,1) \ : \ \mathbb{E}\{\Sigma_k\} \text{ is bounded in } \mathbb{S}^n_{\geq 0}\}.$$

Now, for a given β, if $\gamma_1 \geq \gamma_2$, we have

$$\begin{aligned}\mathbb{E}_{\beta}\{\Sigma^{\gamma_2}_{k+1}\} &= (1-\beta)\mathbb{E}\{h_1(\gamma_2,\Sigma_k)\} + \beta\mathbb{E}\{h_2(\gamma_2,\Sigma_k)\} \\ &\geq (1-\beta)\mathbb{E}\{h_1(\gamma_1,\Sigma_k)\} + \beta\mathbb{E}\{h_2(\gamma_1,\Sigma_k)\} = \mathbb{E}_{\beta}\{\Sigma^{\gamma_1}_{k+1}\},\end{aligned}$$

where the inequality follows from Lemma A.3(ii). Therefore, $\gamma_c^*(\beta)$ can be defined as

$$\gamma_c^*(\beta) := \inf\{\gamma > 0 \ : \ \mathbb{E}\{\Sigma_k\} \text{ is bounded in } \mathbb{S}^n_{\geq 0}\}.$$

This completes the proof. □

The previous theorem shows that under the existence condition of the SMSE, there are critical values that determines boundedness of $\mathbb{E}\{\Sigma_k\}$. Precise characterizations of $\beta_c(\gamma)$ and $\gamma_c^*(\beta)$ are even harder than in the Kalman filtering case due to the worst-case scenario. However, it is still possible to obtain lower and upper bounds for them under the following assumption:

Assumption 2.1. *$h(\gamma, \beta, \Sigma_k)$ is concave in Σ_k for all k.*

This condition guarantees the concavity of the GSRE in the expectation sense. It is shown in Appendix A.2 and Lemma A.4 that there are certain ranges of γ and β for which the above assumption holds. Note that in the rest of this subsection, S_k are F_k are as defined in Result A.1 in Appendix A.3, which constitute respectively lower and upper bounds of $\mathbb{E}\{P_k\}$.

Theorem 2.3. *Suppose that the conditions in Proposition 2.1 and Assumption 2.1 hold, and $Q_0^{-1} \leq DD^T$. Then:*

(i) $\beta_c(\gamma)$ satisfies $\breve{\beta}_c \leq \beta_c(\gamma) \leq \bar{\beta}_c(\gamma)$.

(ii) $\gamma_c^*(\beta)$ satisfies $\breve{\gamma}^* \leq \gamma_c^*(\beta) \leq \bar{\gamma}^*(\beta)$.

(iii) $S_k \leq \breve{\Sigma}_k \leq \mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}_k$ for all k.

(iv) Suppose that $\gamma > \bar{\gamma}^*(\beta)$, $\beta > \bar{\beta}_c(\gamma)$, and γ is finite. Then, $\lim_{k\to\infty} \breve{\Sigma}_k = \breve{\Sigma}^+$ and $\lim_{k\to\infty} \bar{\Sigma}_k = \bar{\Sigma}^+$.

Proof. We first show that $\breve{\Sigma}_k \leq \mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}_k$ for all k. Clearly, $\breve{\Sigma}_0 = \bar{\Sigma}_0 = \mathbb{E}\{\Sigma_0\} = Q_0^{-1}$. Then, by induction, $\breve{\Sigma}_k \leq \mathbb{E}\{\Sigma_k\}$ implies

$$\begin{aligned}\mathbb{E}\{\Sigma_{k+1}\} &\overset{(a)}{=} (1-\beta)\mathbb{E}\{h_1(\gamma, \Sigma_k)\} + \beta\mathbb{E}\{h_2(\gamma, \Sigma_k)\} \\ &\overset{(b)}{\geq} (1-\beta)\mathbb{E}\{h_1(\gamma, \Sigma_k)\} \\ &\overset{(c)}{\geq} (1-\beta)h_1(\gamma, \mathbb{E}\{\Sigma_k\}) \overset{(d)}{\geq} (1-\beta)h_1(\gamma, \breve{\Sigma}_k) = \breve{\Sigma}_{k+1},\end{aligned}$$

where (a) follows from the law of iterated expectations, (b) is due to Lemma A.3(v) in Appendix A.2, (c) follows from Jensen's inequality due to Lemma A.3(iii), and (d) is due to the induction argument and Lemma A.3(viii). Similarly, $S_k \leq \breve{\Sigma}_k$ can be shown by induction and Lemma A.3(v).

For the second part, $\mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}_k$ implies

$$\mathbb{E}\{\Sigma_{k+1}\} \overset{(e)}{=} \mathbb{E}\{h(\gamma, \beta, \Sigma_k)\} \overset{(f)}{\leq} h(\gamma, \beta, \mathbb{E}\{\Sigma_k\}) \overset{(g)}{\leq} h(\gamma, \beta, \bar{\Sigma}_k) = \bar{\Sigma}_{k+1},$$

where (e) is due to the law of iterated expectations, (f) follows due to the assumption and from Jensen's inequality, and (g) is obtained from the induction argument and Lemma A.3(viii).

Now, if $\gamma < \breve{\gamma}^*$ and $\beta < \breve{\beta}_c$, then since $\sqrt{1-\beta}A$ is not stable, $\{S_k\} \to \infty$ as $k \to \infty$ due to Theorem 3 in [22]. Therefore, $\{\breve{\Sigma}_k\} \to \infty$ as $k \to \infty$ and $\mathbb{E}\{\Sigma_k\} \to \infty$ as $k \to \infty$. This implies that $\beta_c(\gamma) \geq \breve{\beta}_c$ and $\gamma_c^*(\beta) \geq \breve{\gamma}^*$. Let us consider the second case. By Proposition 2.3 and the previous inequality, we have that $\{\bar{\Sigma}_k\} \to \bar{\Sigma}^+$ as $k \to \infty$, and $\bar{\Sigma}^+ \geq \mathbb{E}\{\Sigma_k\}$ for all k. Since $\bar{\Sigma}^+ \geq \breve{\Sigma}_k$, we have convergence of the lower bound due to Proposition 2.2. Hence, we have $\beta_c(\gamma) \leq \bar{\beta}_c(\gamma)$ and $\gamma_c^*(\beta) \leq \bar{\gamma}^*(\beta)$. This completes the proof of the theorem. □

The following series of discussions provides a comparison of Theorem 2.3 with Result A.1 in Appendix A.3.

Remark 2.2. *We have $S_k \leq \breve{\Sigma}_k$ and $F_k \leq \bar{\Sigma}_k$ due to Lemma A.3(v) and (ii). Moreover, $\breve{\Sigma}_k \to S_k$ and $\bar{\Sigma}_k \to F_k$ as $\gamma \to \infty$ from Lemma A.3(vi) and (vii), respectively. This shows that since $P_k \leq \Sigma_k$ a.s. for all k, $S_k \leq \mathbb{E}\{P_k\} \leq \mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}_k$, and the lower and upper bounds converge to the deterministic values in view of Theorem 2.3. This implies that given γ and β, if $\mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}^+$, then $\mathbb{E}\{P_k\} \leq \bar{\Sigma}^+$, that is, the Kalman filter is also stable in the expectation sense.*

Remark 2.3. *Remark 2.2 implies that given γ and β, if we have $\mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}^+$, then $\beta \geq \bar{\lambda}$, where $\bar{\lambda}$ is the upper critical value of the Kalman filter defined in Appendix A.3. Therefore, $\beta \geq \bar{\lambda}$ is a* necessary condition *for boundedness of $\mathbb{E}\{\Sigma_k\}$. On the other hand, for the Kalman filtering case, $\beta \geq \bar{\lambda}$ is a sufficient condition for the boundedness of $\mathbb{E}\{P_k\}$ (see Appendix A.3).*

Remark 2.4. *As $\gamma \to \infty$, $\mathbb{E}\{\Sigma_k\} \to \mathbb{E}\{P_k\}$. Moreover, when $\gamma \to \infty$, its boundedness condition is equivalent to that in Result A.1, since $\breve{\beta}_c = \breve{\lambda}$ and $\bar{\beta}_c(\gamma) \to \bar{\lambda}$ in view of Proposition 2.4. In this case, Assumption 2.1 holds for any β as discussed in Appendix A.2.*

Remark 2.5. *Even if C is invertible, there is no reason that $\beta_c(\gamma) = \breve{\beta}_c$ unless γ is sufficiently large. This is because the SMSE considers the worst-case scenario that results in increasing $\bar{\beta}(\gamma)$ depending on γ due to Proposition 2.4(ii). This fact will be demonstrated in Section 2.5. This implies that the statement of Result A.1(iii) is not valid for the minimax estimation problem.*

2.4.2 Weak Convergence of $\{\|\Sigma_k\|\}$

This subsection presents a result on weak convergence of the norm of the sequence generated by the GSRE, as captured in the following theorem.

Theorem 2.4. *Suppose that the condition in Proposition 2.1 holds, $\beta > 0$, and $DD^T > 0$. Then, $\{\|\Sigma_k\|\}$[3] converges weakly to a unique stationary distribution from any initial condition $\Sigma_0 = Q_0^{-1}$.*

Proof. Under the existence condition, since $P_k \leq \Sigma_k$ a.s. for all k, for any $M \geq 0$ we have

$$\begin{aligned}\mathbb{P}(\|P_k\| \leq M) &\geq \mathbb{P}(\|P_k\| \leq M,\ \|\Sigma_k\| \leq M)\\ &= \mathbb{P}(\|P_k\| \leq M \mid \|\Sigma_k\| \leq M)\mathbb{P}(\|\Sigma_k\| \leq M) = \mathbb{P}(\|\Sigma_k\| \leq M),\end{aligned}$$

where $\mathbb{P}(\cdot|\cdot)$ is conditional probability. We let $F_{\Sigma_k}(M) := \mathbb{P}(\|\Sigma_k\| \leq M)$ and $F_{P_k}(M) := \mathbb{P}(\|P_k\| \leq M)$. Clearly, $\{F_{P_k}(M)\}$ is a sequence of distribution functions of the norm of the matrices generated by the SRE in (A.14) (or (A.15)) for all k. In the same vein, $\{F_{\Sigma_k}(M)\}$ is a sequence of distribution functions of the norm of the matrices generated by the GSRE in (2.7) for all k. Also, by definition, $F_{\Sigma_k}(M) \leq F_{P_k}(M)$ for all k. Then, if $\{F_{\Sigma_k}(M)\}$ converges to some distribution function at all continuity points M of its limit, we have the weak convergence [108].

For any $\beta > 0$, $\{\|P_k\|\}$ converges weakly to a unique stationary distribution due to Theorem 3.1 in [54]. We can also equivalently say that for any $\beta > 0$, the sequence of distribution functions $\{F_{P_k}(M)\}$ converges to $F_P(M)$ as $k \to \infty$ at all continuity points of M of $F_P(M)$, where $F_P(M)$ is a distribution function [108]. Therefore, $\lim_{k\to\infty} F_{\Sigma_k}(M) \leq F_P(M)$.

Now, we use the following fact that follows from the convergence of sequences of real numbers [109]: a sequence of real numbers, say $\{x_k\}$, converges to x if and only if every subsequence of $\{x_k\}$ converges to x.

By using Helly's selection theorem [108], for the sequence of distribution functions $\{F_{\Sigma_k}(M)\}$, there is a subsequence, $\{F_{\Sigma_{k(l)}}(M)\}$, and a right continuous nondecreasing function $F_\Sigma(M)$ so that $\lim_{l\to\infty} F_{\Sigma_{k(l)}}(M) = F_\Sigma(M)$ at all continuity points M of $F_\Sigma(M)$.

We claim that $F_\Sigma(M)$ is a distribution function. To show this, due to [108, Theorem 2.6], it suffices to show that the sequence $\{F_{\Sigma_k}(M)\}$ is tight, i.e.,

[3]Note that $\|X\|$ is the induced 2-norm of the matrix X.

for all $\epsilon > 0$, there is an M_ϵ so that $\limsup_{k\to\infty} 1 - F_{\Sigma_k}(M_\epsilon) + F_{\Sigma_k}(-M_\epsilon) \leq \epsilon$. The tightness can also be written as for each $\epsilon > 0$, there exist $M_\epsilon^{(1)}$ and $M_\epsilon^{(2)}$ such that $F_{\Sigma_k}(M_\epsilon^{(1)}) < \epsilon$ and $F_{\Sigma_k}(M_\epsilon^{(2)}) > 1 - \epsilon$, for all k [110].

Note that the sequence $\{F_{P_k}(M)\}$ is tight in view of [108, Theorem 2.6]. Then, for each $\epsilon > 0$, there exists $M_\epsilon^{(1)}$ such that $F_{\Sigma_k}(M_\epsilon^{(1)}) \leq F_{P_k}(M_\epsilon^{(1)}) < \epsilon$ for all k. Moreover, for all k, we have

$$1 - F_{P_k}(M) \leq 1 - F_{\Sigma_k}(M) \Leftrightarrow \mathbb{P}(\|P_k\| > M) \leq \mathbb{P}(\|\Sigma_k\| > M),$$

and $\lim_{M\to\infty} \mathbb{P}(\|P_k\| > M) = \lim_{M\to\infty} \mathbb{P}(\|\Sigma_k\| > M) = 0$. This implies that for each $\epsilon > 0$, there exists $M_\epsilon^{(2)}$ such that $\mathbb{P}(\|\Sigma_k\| > M_\epsilon^{(2)}) < \epsilon$ for all k, which leads to $1 - \mathbb{P}(\|\Sigma_k\| > M_\epsilon^{(2)}) = F_{\Sigma_k}(M_\epsilon^{(2)}) > 1 - \epsilon$. This proves the claim.

We have shown that for the sequence of distribution functions $\{F_{\Sigma_k}(M)\}$, there exists a convergent subsequence $\{F_{\Sigma_{k(l)}}(M)\}$, where its limit, $F_\Sigma(M)$, is also a distribution function. To complete the proof, we need to show that every subsequence converges to $F_\Sigma(M)$.

Consider any arbitrary subsequence $\{F_{\Sigma_{m(l)}}(M)\}$ of $\{F_{\Sigma_k}(M)\}$. Note that $\{F_{\Sigma_{m(l)}}(M)\}$ and $\{F_{\Sigma_{k(l)}}(M)\}$ are different subsequences of $\{F_{\Sigma_k}(M)\}$ indexed by l. Moreover, since $\{F_{P_k}(M)\}$ is a convergent sequence, by using the above fact, the subsequences, $\{F_{P_{k(l)}}(M)\}$ and $\{F_{P_{m(l)}}(M)\}$, also converge to $F_P(M)$ as $l \to \infty$.

For each $\epsilon \geq 0$, there exist l_1 and l_2 such that

$$\begin{aligned}
|F_{\Sigma_{k(l)}}(M) - F_\Sigma(M)| &\leq \frac{\epsilon}{2},\ \forall l \geq l_1 \\
|F_{\Sigma_{m(l)}}(M) - F_{\Sigma_{k(l)}}(M)| \leq |F_{P_{k(l)}}(M) - F_P(M)| &\leq \frac{\epsilon}{2},\ \forall l \geq l_2,
\end{aligned}$$

where we made use of the fact that $|F_{\Sigma_{m(l)}}(M) - F_{\Sigma_{k(l)}}(M)| \leq 1$ and $F_{\Sigma_k}(M) \leq F_{P_k}(M)$ for all k.

Let $l' := \max\{l_1, l_2\}$. Then, for any $l \geq l'$,

$$\begin{aligned}
|F_{\Sigma_{m(l)}}(M) - F_\Sigma(M)| &\leq |F_{\Sigma_{k(l)}}(M) - F_\Sigma(M)| + |F_{\Sigma_{m(l)}}(M) - F_{\Sigma_{k(l)}}(M)| \\
&\leq |F_{\Sigma_{k(l)}}(M) - F_\Sigma(M)| + |F_{P_{k(l)}}(M) - F_P(M)| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since ϵ was arbitrary, $\{F_{\Sigma_{m(l)}}(M)\}$ also converges to $F_\Sigma(M)$ as $l \to \infty$. Note that this holds for any arbitrary subsequence. Hence, every subsequence of $\{F_{\Sigma_k}(M)\}$ has the same limit, which proves the theorem. $\square$

Remark 2.6. *The weak convergence of* $\{\|P_k\|\}$ *to a unique stationary distribution was discussed in [52] and [54]. While [52] considered the SRE in (A.14) (or (A.15)) as a random walk and then proved weak convergence by using the mean contraction property when* A *is non-singular, [54] showed the same result by modeling the SRE as an order-preserving and strongly sublinear random dynamical system. On the other hand, in this chapter, under the existence condition, we have proven the weak convergence of* $\{\|\Sigma_k\|\}$ *by showing the convergence of the sequence of distribution functions.*

Remark 2.7. *When* $\gamma \to \infty$*, the stationary distribution in (i) converges to that in [54, Theorem 3.1]. Moreover, Theorem 2.4 also holds when* $\|\Sigma_k\|^2 = \mathrm{Tr}(\Sigma_k^T \Sigma_k)$ *where* $\mathrm{Tr}(\cdot)$ *is the trace operator.*

2.5 Simulations: Asymptotic Analysis and Estimation Performance

In this section, we present simulation results on the asymptotic analysis of the GSRE discussed in Section 2.4.1, and compare the estimation performance with the H^∞ MJLE in [58].

2.5.1 SMSE and Kalman Filter: Lower and Upper Bounds

We first analyze the lower and upper bounds on $\mathbb{E}\{\Sigma_k\}$ and $\mathbb{E}\{P_k\}$. To allow for a comparison, we use the same linear system model as in [22]. We first consider the scalar case where $A = -1.25$, $C = 1$, $D = 1$, $V = 2.5$, and $Q = 1$. For the Kalman filter case, $\lambda_c = 0.36$. Moreover, it can be calculated that $\bar{\gamma}^*(1) = 1.585$. Figure 2.1(a) shows a plot of the convergence region of (2.12). This plot is obtained by using the following approach:

(S.1) Fix $\beta = 1$ and take a sufficiently large value of $\gamma > 0$.

(S.2) Obtain the solution, $\bar{\Sigma} \in \mathbb{S}^n_{>0}$, of the MGARE in (2.14), and check the condition $\rho(\bar{\Sigma} Q) < \gamma^2$.

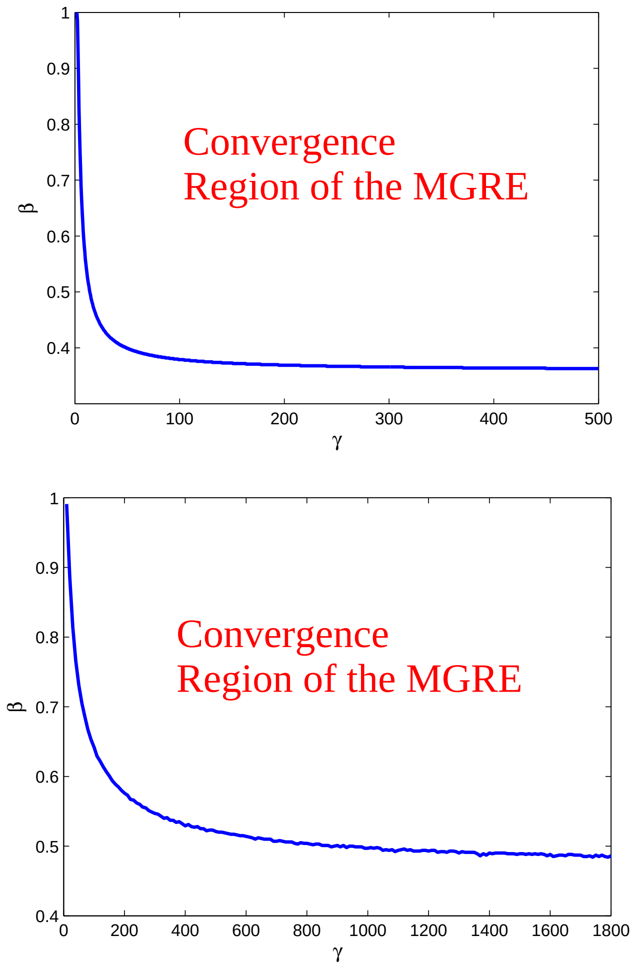


Figure 2.1: Convergence region of (2.12) (top: scalar case, bottom: matrix case).

(S.3) If the condition holds, decrease β and then go to (S.2). Otherwise, it is the critical value of β for that γ; go to the next step.

(S.4) Decrease γ and fix $\beta = 1$. Go to (S.2).

As can be seen, as $\gamma \to \infty$, $\bar{\beta}_c(\gamma) \to 0.36 = \breve{\beta}_c$. Figure 2.2 shows a plot of the steady state upper and lower bounds on $\mathbb{E}\{\Sigma_k\}$ versus β. In this simulation, although C is invertible, since $\bar{\beta}_c(\gamma) = 0.36$ only when γ is sufficiently large, $\breve{\beta}_c$ cannot be the critical value for boundedness of $\mathbb{E}\{\Sigma_k\}$, provided that the existence condition in Proposition 2.1 and Assumption 2.1 hold. That is, $\beta \geq \breve{\beta}_c$ is necessary but not sufficient for boundedness of $\mathbb{E}\{\Sigma_k\}$. This result is a direct consequence of Proposition 2.4 and is due to the worst-case approach.

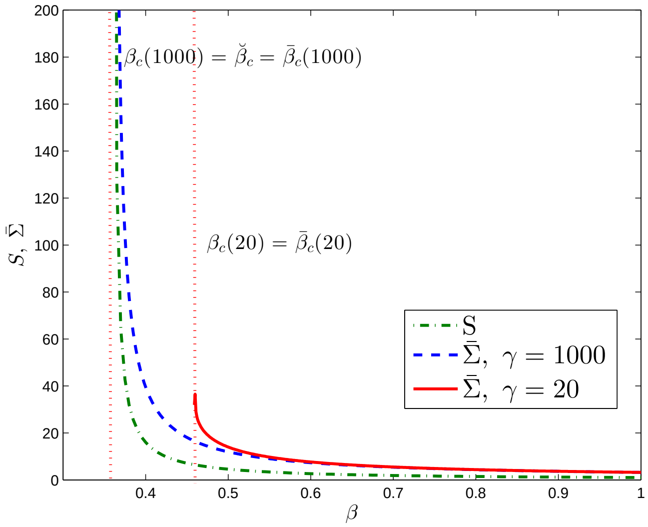


Figure 2.2: Transition to instability in the scalar case.

We now discuss the two-dimensional case, where

$$A = \begin{pmatrix} 1.25 & 0 \\ 1 & 1.1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 1 \end{pmatrix}, \ D = \begin{pmatrix} \sqrt{20} & 0 \\ 0 & \sqrt{20} \end{pmatrix}, \ V = 2.5, \ Q = 20.$$

Figure 2.1(b) shows the convergence region of (2.12). The plot is also obtained by using the above approach. Note that $\bar{\beta}_c(\gamma) \to \bar{\lambda}$ as $\gamma \to \infty$. The plot of the lower and upper bounds on $\mathbb{E}\{\Sigma_k\}$ for the two-dimensional case is shown in Fig. 2.3. Note that the critical transition for boundedness of $\mathbb{E}\{\Sigma_k\}$ depends on the level of the disturbance attenuation parameter γ. Moreover, as mentioned, $\beta \geq \bar{\lambda}$ is necessary for boundedness of $\mathbb{E}\{\Sigma_k\}$.

2.5.2 SMSE versus H^∞ MJLE

We compare estimation performance of the SMSE with that of the H^∞ MJLE in [58]. Figure 2.4 shows the existence region of the SMSE for the same scalar system as in Section 2.5.1. We obtained this plot by Monte Carlo simulations with 1000 samples of the GSRE.

Now, to use the H^∞ MJLE theory in [58], the transition matrix S is given

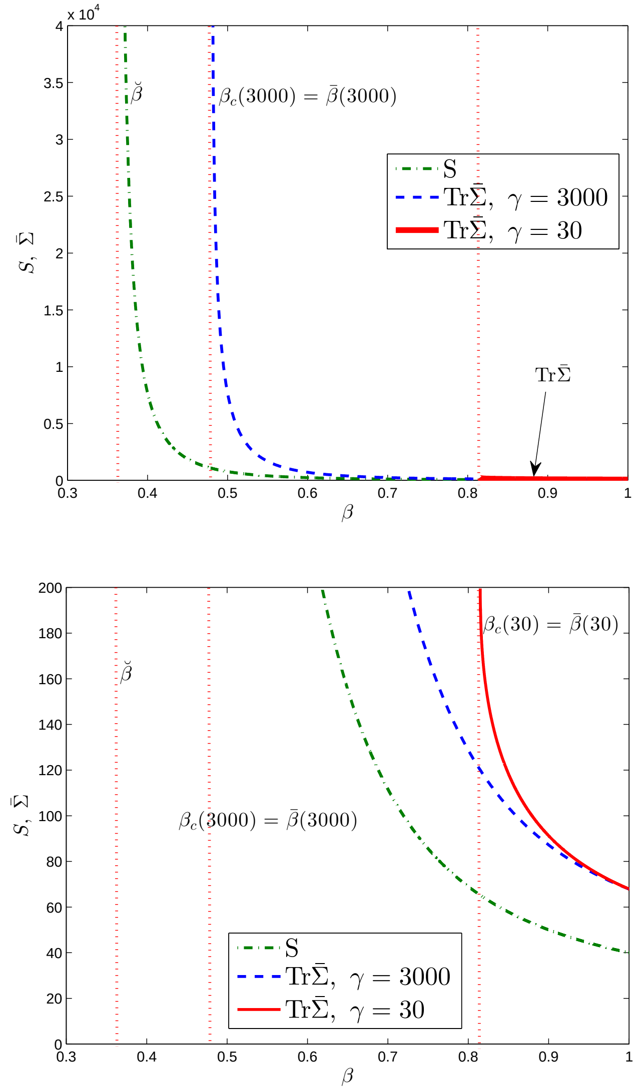


Figure 2.3: Transition to instability in the matrix case.

by

$$S = \begin{pmatrix} 1-\beta & \beta \\ 1-\beta & \beta \end{pmatrix}.$$

It is easy to show that the i.i.d. Bernoulli process $\{\beta_k\}$ with $\mathbb{P}(\beta_k = 1) = \beta$ is identical to the Markov process $\{\beta_k\}$ with S. Also, the corresponding optimum disturbance attenuation level when $\beta = 0.7$ is 7.5.

Figure 2.5 shows $\mathbb{E}\{\Sigma_k\}$ and $\bar{\Sigma}_k$ when $\gamma = 40$ and $\beta = 0.7$. Note that we have $\mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}_k$ for all k; therefore, Assumption 2.1 holds. Figure 2.6 shows

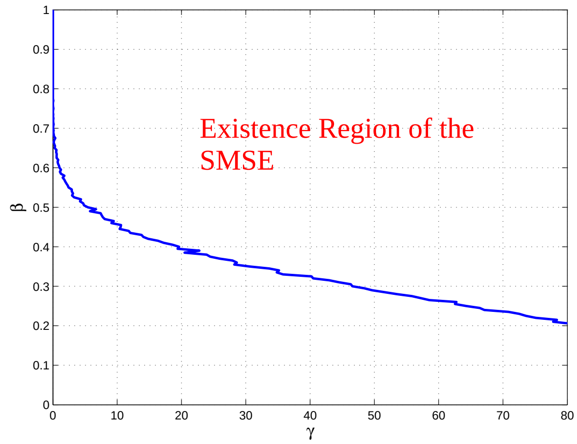


Figure 2.4: Existence region of the SMSE of the scalar system. The plot is obtained by Monte Carlo simulations with 1000 samples of (2.7).

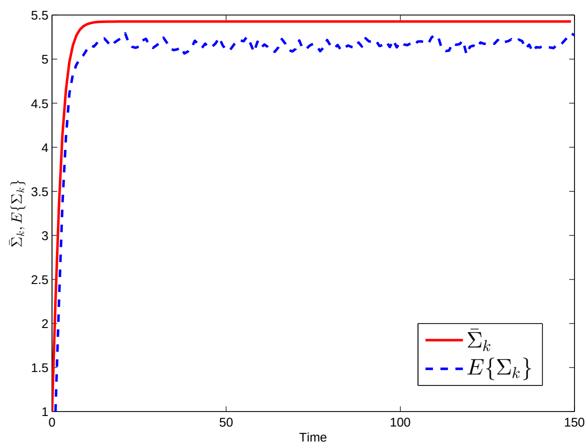


Figure 2.5: $\mathbb{E}\{\Sigma_k\}$ and $\bar{\Sigma}_k$ of the scalar system. The plot of $\mathbb{E}\{\Sigma_k\}$ is obtained by averaging 10,000 Monte Carlo simulations.

the mean-square estimation error of the SMSE and the H^∞ MJLE when $\beta = 0.7$, and $w_k = -v_k = 10$. Each curve is obtained by averaging 10,000 Monte Carlo simulations. We use $\gamma = 40$ for the SMSE. Due to Figures 2.4 and 2.1(a), and Theorem 2.3, such a choice guarantees existence of the SMSE and $\mathbb{E}\{\Sigma_k\} \leq \bar{\Sigma}$. Note that although γ for the SMSE is larger than that of the H^∞ MJLE (that is 7.5), the better estimation error is achieved by the SMSE. This is expected, since the SMSE uses $\{\beta_k\}$ that corresponds to the entire information on the measurement arrivals, whereas the H^∞ MJLE uses only the instantaneous information β_k. Finally, the SMSE with the static

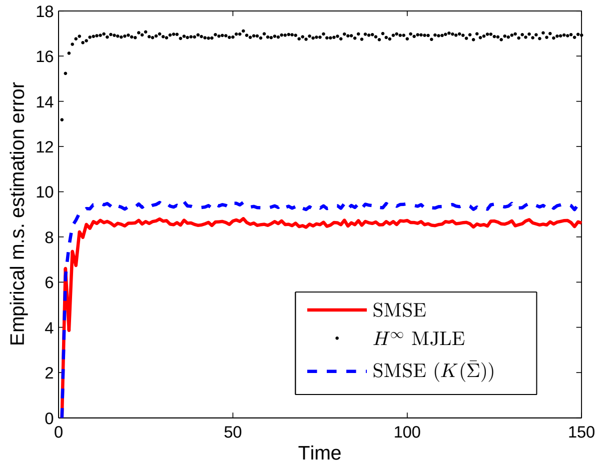


Figure 2.6: Empirical r.m.s. estimation error when $w_k = 10$ and $v_k = -10$. This plot is obtained by averaging 10,000 Monte Carlo simulations.

gain obtained from $\bar{\Sigma}$ also shows better estimation performance than the H^∞ MJLE, since $\bar{\Sigma}$ constitutes the tight upper bound for $\mathbb{E}\{\Sigma_k\}$.

2.6 Conclusions

In this chapter, we have considered the problem of minimax estimation with intermittent observations. Unlike the previous work, we have considered the situation when the sensor noise and the disturbance are not stochastic processes, but are treated as adversarial inputs. As such, the problem has been formulated within the framework of stochastic zero-sum dynamic games.

We have obtained the stochastic minimax state estimator (SMSE) and the associated generalized stochastic Riccati equation (GSRE). We have identified a threshold-type existence condition for the SMSE as a function of the disturbance attenuation parameter, γ, and also one for the GSRE, which implicitly depends on the measurement arrival rate β. We have shown that when the existence condition holds for a particular γ, the SMSE achieves the disturbance attenuation level corresponding to that γ, and if $\gamma \to \infty$, the SMSE converges to the Kalman filter with intermittent observations.

Two different asymptotic behaviors of the GSRE have been discussed. Specifically, we have shown boundedness of the sequence generated by the GSRE in the expectation sense, and weak convergence of the norm of that sequence to a unique stationary distribution. These two results indicate that stability of the SMSE implies stability of the Kalman filter with the same unreliable communication channels. This is a consequence of the fact that the SMSE should be more conservative than the Kalman filter, since the former deals with arbitrary disturbances under the worst-case scenario.

CHAPTER 3

MINIMAX CONTROL OVER UNRELIABLE COMMUNICATION CHANNELS

3.1 Introduction

In this chapter, we study a minimax control problem over unreliable communication channels. Unlike the previous work on control over unreliable communication channels, as mentioned in Chapter 1, this chapter considers the case when the disturbance and the sensor noise in dynamical systems are arbitrary and controlled by adversaries, instead of being stochastic with *a priori* specified statistics. We consider two different scenarios for unreliable communication channels: the TCP- and the UDP-cases, as depicted in Figs. 1.5 and 1.6, respectively. Both communication channels are assumed to be temporally uncorrelated, which are modeled as two independent and identically distributed Bernoulli processes. These two different problems are formulated within the framework of stochastic zero-sum dynamic games, which enables us to develop worst-case (H^∞) controllers under TCP- and UDP-like information structures.

We first consider the TCP-case. Due to its acknowledgment nature, we are able to apply the certainty equivalence principle developed in [106] and [107], where the deterministic H^∞ optimal control was analyzed through three steps. By following these steps, we obtain a class of output feedback minimax controllers in terms of the H^∞ disturbance attenuation parameter, say γ, and the control and measurement loss rates, where γ is a parameter that measures robustness of the system against arbitrary disturbances as in standard H^∞ control [106, 111]. For the TCP-case, the minimax controller obtained is dependent on the acknowledged control packet loss information, which differs from the existing mode-dependent H^∞ controller for MJLSs in the literature as discussed in Chapter 1.

Specifically, the main results for the TCP-case are as follows:

(i) The existence of a minimax controller is dependent on γ, and the loss rates.

(ii) For given loss rates and $\gamma > 0$, if all the existence conditions are satisfied, then γ is the attenuation level of the corresponding minimax controller.

(iii) The critical values of the control and measurement loss rates for closed-loop system stability and performance are functions of γ.

(iv) There is no separation between control and estimation.

(v) As $\gamma \to \infty$, the parametrized (in γ) minimax control system converges to the corresponding LQG control system in [19] and [21].

Item (ii) implies that if γ exists and is finite, then the corresponding admissible minimax controller achieves the disturbance attenuation level γ for an arbitrary disturbance, that is, the H^∞ norm of the closed-loop system is bounded above by γ [106]. As for item (v), the limiting behavior in terms of γ implies that in view of item (ii), the disturbance does not play any role, since it is infinitely penalized; hence, the limiting behavior of the corresponding minimax controller collapses to the LQG controller as in the standard case discussed in [106, 111].

For the UDP-case, we consider the scenario when there is no measurement noise, which is a counterpart of the LQG problem discussed in [19]. We show that due to the absence of acknowledgments regarding control packet losses, there is dual effect between control and estimation, but the corresponding minimax controller parametrized by γ is linear in the measurement. Such a dual effect problem did not arise in the H^∞ control problem within the MJLS framework, since as already mentioned, the latter has access to the current mode of the Markov chain. We provide the (different) existence condition for the corresponding problem in terms of γ and control and measurement loss rates. We also provide explicit expressions on the H^∞ optimum disturbance attenuation parameter and the critical values for mean-square stability and performance of the closed-loop system. Moreover, we show that when $\gamma \to \infty$, the minimax control system collapses to the corresponding LQG system in [19]. Finally, from simulation results, we show that the stability and performance regions for the UDP-case are more stringent than those of the TCP-case due to lack of acknowledgments.

Organization

The organization of the chapter is as follows. The problem formulation is stated in Section 3.2. Sections 3.3, 3.4, and 3.5 are for the TCP-case, which consider problems of state feedback minimax control, minimax estimation with intermittent observations (for which complete development can be found in Chapter 2), and the H^∞ synthesis problem, respectively. A special case of the UDP problem is studied in Section 3.6. Section 3.7 provides numerical examples. We end the chapter with the concluding remarks of Section 3.8.

3.2 Problem Formulation

We consider the following linear dynamical system:

$$x_{k+1} = Ax_k + \alpha_k B u_k + D w_k \tag{3.1a}$$

$$y_k = \beta_k C x_k + E w_k, \tag{3.1b}$$

where $x_k \in \mathbb{R}^n$ is the state; $u_k \in \mathbf{U} \subset \mathbb{R}^m$ is the control (actuator); $w_k \in \mathbf{W} \subset \mathbb{R}^p$ is the disturbance input as well as the measurement noise; $y_k \in \mathbf{Y} \subset \mathbb{R}^l$ is the sensor output; and A, B, C, D, E are time-invariant matrices with appropriate dimensions. In (3.1), $\{w_k\}$ is a square-summable sequence, which is not necessarily stochastic. We further assume the following decompositions:

$$w_k = \begin{pmatrix} \bar{w}_k \\ v_k \end{pmatrix}, \quad D = \begin{pmatrix} \bar{D} & 0 \end{pmatrix}, \; E = \begin{pmatrix} 0 & \bar{E} \end{pmatrix}, \; V = \bar{E}\bar{E}^T > 0,$$

where $\bar{E}$ is square and non-singular. Finally, we assume that the communication network is temporally uncorrelated, that is, $\{\alpha_k\}$ and $\{\beta_k\}$ in (3.1) are independent and identically distributed (i.i.d.) Bernoulli processes with $\mathbb{P}(\alpha_k = 1) = \alpha$ and $\mathbb{P}(\beta_k = 1) = \beta$, respectively. We denote the variance of α_k by $\bar{\alpha} := \alpha(1-\alpha)$.

The TCP-like information that is available to the controller is defined by

$$\begin{cases} \mathcal{I}_0 & := \{y_0, \beta_0\} \\ \mathcal{I}_k & := \{y_{0:k}, \alpha_{0:k-1}, \beta_{0:k}\}, \; k \geq 1. \end{cases} \tag{3.2}$$

The UDP-like information is defined by

$$\begin{cases} \mathcal{G}_0 & := \{y_0, \beta_0\} \\ \mathcal{G}_k & := \{y_{0:k}, \beta_{0:k}\},\ k \geq 1. \end{cases} \tag{3.3}$$

Note that the major difference between (3.2) and (3.3) is that in (3.2), the acknowledgment signal, $\alpha_{0:k-1}$, is included, by which, as expected from the LQG case in [19], the controller under (3.2) will provide better stability and performance.

A convention we adopt in this chapter is one of zero-input strategy. That is, the actuator does not do anything when there are control packet losses. It was shown in [112] that for the disturbance free case ($w_k \equiv 0$), using the one-step previous control packet in order to compensate for the current packet loss does not necessarily lead to better performance.

Let $\mathcal{U}$ and $\mathcal{W}$ be the appropriate spaces of control and disturbance policies, respectively. We define control and disturbance policies, $\mu \in \mathcal{U}$ and $\nu \in \mathcal{W}$, that consist of sequences of functions:

$$\mu = \{\mu_0, \mu_1, ..., \mu_{N-1}\},\ \nu = \{\nu_0, \nu_1, ..., \nu_{N-1}\},$$

where μ_k and ν_k are Borel measurable functions which map the information set (3.2) or (3.3) into the control and disturbance spaces of $\mathbb{R}^m$ and $\mathbb{R}^p$, respectively. Note that in the spirit of the worst-case approach, the disturbance is assumed to know everything the controller does.

Now, our main objective in this chapter is to obtain output feedback controllers over TCP- and UDP-networks, which minimize the following cost function:

$$\ll \mathcal{T}_\mu^N \gg := \sup_{(x_0, w_{0:N-1})} \frac{J^N(\mu,\nu)^{1/2}}{\mathbb{E}\Big\{|x_0|^2_{Q_0} + \sum_{k=0}^{N-1} |w_k|^2\Big\}^{1/2}}, \tag{3.4}$$

where

$$J^N(\mu,\nu) = \mathbb{E}\Big\{|x_N|^2_{Q_N} + \sum_{k=0}^{N-1} |x_k|^2_Q + \alpha_k |u_k|^2_R\Big\},$$

where $Q, Q_N \geq 0,\ R, Q_0 > 0$, and μ and ν are the control and disturbance policies as introduced earlier. Note that the control u_k incurs the additional cost only if it is applied to the plant. This can be viewed as an H^∞ optimal control problem [106]. It is worth noting that if α_k is included in (3.2), then the problem can be studied in the framework of Markov jump linear systems, and the optimal controller can then be obtained directly from [36].

Associated with the system (3.1), we introduce the zero-sum dynamic game that is parametrized by the disturbance attenuation parameter, $\gamma > 0$:

$$J_\gamma^N(\mu, \nu) = \mathbb{E}\Big\{|x_N|^2_{Q_N} - \gamma^2|x_0|^2_{Q_0} + \sum_{k=0}^{N-1} |x_k|^2_Q + \alpha_k|u_k|^2_R - \gamma^2|w_k|^2\Big\}, \quad (3.5)$$

subject to system (3.1a) and the measurement equation (3.1b).

Now, in view of (3.4) and (3.5), our main objective in this chapter can be rephrased as one of obtaining a controller for (3.1) under the specified information structure such that it minimizes the cost function (3.5) while the disturbance maximizes the same cost function. In other words, we need to characterize a saddle point,[1] say (μ_γ, ν_γ), for the zero-sum dynamic game (3.5) in terms of γ.

As in standard H^∞ control [106], the existence of saddle-point solutions for (3.5) will be dependent on the value of γ. Therefore, we need to characterize the smallest value of γ, say γ^*, above which the saddle-point solutions exist. Then, by definition, for any $\gamma > \gamma^*$, the saddle point, (μ_γ, ν_γ), exists, provided that γ^* is finite. Moreover, for any finite $\gamma > \gamma^*$, μ_γ is a minimax controller that leads to finite upper value for the zero-sum dynamic game in (3.5), and achieves the performance level of γ for (3.4), i.e., under μ_γ, $\ll \mathcal{T}^N_{\mu_\gamma} \gg \leq \gamma$.

After characterizing a class of minimax controllers for the TCP- and UDP-cases, the next goal is to examine such controllers with respect to the communication channel conditions. Specifically, given the controllers, we need to obtain the smallest values of α and β, say α_c and β_c, for the closed-loop system stability and performance. Obviously, α_c and β_c are functions of γ, and γ^* is a function of α and β.

In what follows, in Sections 3.3-3.5, we obtain a class of output feedback minimax controllers for the TCP-case. Toward that end, we apply the cer-

[1]See [106] and [61] for the definition of saddle point for zero-sum dynamic games. Normally, in going from (3.4) to (3.5), one would be looking for the minimax solution of (3.5), but as in [106], one could instead look for the saddle-point solution, without any loss of generality.

tainty equivalence principle discussed in Appendix B.2, in view of which the corresponding zero-sum dynamic game can be analyzed through three steps discussed in Appendix B.2. Note that the certainty equivalence principle was originally developed by [106] and [107] for the deterministic (no packet drops) H^∞ control problem, and the results presented in Appendix B.2 can be regarded as the *certainty equivalence principle* of the H^∞ control problem for the TCP-case.

In Section 3.6, we obtain a (different) class of output feedback minimax controllers for the UDP-case. We consider a special case of this problem, where there is no measurement noise in (3.1b). As discussed in Section 3.6, the general minimax control problem for the UDP-case is hard, since there is no acknowledgment of control packet losses. In view of the certainty equivalence principle in Appendix B.2, this is because part (b) of the certainty equivalence principle cannot be applied to the UDP-case, which is shown in Section 3.4.

3.3 State Feedback Minimax Control over the TCP-Network

This section addresses part (a) of the certainty equivalence principle discussed in Appendix B.2. In particular, we obtain a state feedback minimax controller over the TCP-network.

3.3.1 Finite-Horizon Case

Lemma 3.1. *Consider the zero-sum dynamic game in (3.5) with a fixed $\gamma > 0$ and $\alpha \in [0,1]$. Then:*

(i) There exists a unique state feedback saddle-point solution if and only if

$$\rho(D^T Z_{k+1} D) < \gamma^2, \text{ for all } k \in [0, N-1], \tag{3.6}$$

where Z_k is generated by the generalized Riccati equation (GRE):

$$Z_k = Q + P_{u_k}^T(\alpha R + \bar{\alpha} B^T Z_{k+1} B) P_{u_k} - \gamma^2 P_{w_k}^T P_{w_k} + H_k^T Z_{k+1} H_k, \tag{3.7}$$

where $Z_N = Q_N$, and

$$H_k = A - \alpha B P_{u_k} + D P_{w_k} \tag{3.8a}$$

$$P_{u_k} = (R + B^T(I + \alpha Z_{k+1} D M_k^{-1} D^T) Z_{k+1} B)^{-1} \times B^T(I + Z_{k+1} D M_k^{-1} D^T) Z_{k+1} A \tag{3.8b}$$

$$P_{w_k} = (\gamma^2 I - D^T(I - \alpha Z_{k+1} B L_k^{-1} B^T) Z_{k+1} D)^{-1} \times D^T(I - \alpha Z_{k+1} B L_k^{-1} B^T) Z_{k+1} A \tag{3.8c}$$

$$M_k = \gamma^2 I - D^T Z_{k+1} D \tag{3.8d}$$

$$L_k = R + B^T Z_{k+1} B. \tag{3.8e}$$

(ii) The feedback saddle-point policies, $(\mu_\gamma^*, \nu_\gamma^*)$, can be written as

$$u_k^* = \mu_k^*(\mathcal{I}_k) = -P_{u_k} x_k \tag{3.9}$$

$$w_k^* = \nu_k^*(\mathcal{I}_k) = P_{w_k} x_k, \ \ k \in [0, N-1]. \tag{3.10}$$

(iii) If M_k has a negative eigenvalue for some k, then the zero-sum dynamic game does not admit a saddle point and the upper value of the game becomes unbounded.

Proof. To prove parts (i) and (ii), we need to employ dynamic programming with the following value function $V_k(x_k) = \mathbb{E}\{x_k^T Z_k x_k | \mathcal{I}_k\}$, where $Z_k \geq 0$ is given in (3.7) with $Z_N = Q_N$ [61]. Now, by induction, suppose the claim is true for $k+1$. That is, $V_{k+1}(x_{k+1})$ is the saddle-point value of the static zero-sum game at $k+1$ under (3.6). Then, since the information structure of the TCP-network is nested for all k, the cost-to-go at k can be written as

$$V_k(x_k) = \min_{u_k} \max_{w_k} \mathbb{E}\Big\{ h_k(x,u,w) + V_{k+1}(x_{k+1}) \Big| \mathcal{I}_k \Big\} \tag{3.11}$$

$$= \max_{w_k} \min_{u_k} \mathbb{E}\Big\{ h_k(x,u,w) + V_{k+1}(x_{k+1}) \Big| \mathcal{I}_k \Big\}, \tag{3.12}$$

where $h_k(x,u,w) := |x_k|_Q^2 + \alpha_k |u_k|_R^2 - \gamma^2 |w_k|^2$. Under (3.6), the static zero-sum game above is strictly convex in u_k and concave in w_k; hence there is a unique pair of minimizer and maximizer, which can be written as

$$u_k^* = -(R + B^T Z_{k+1} B)^{-1} B^T Z_{k+1} (A x_k + D w_k^*) =: \varphi_{1,k}(x_k, w_k^*)$$

$$w_k^* = (\gamma^2 I - D^T Z_{k+1} D)^{-1} D^T Z_{k+1} (A x_k + \alpha B u_k^*) =: \varphi_{2,k}(x_k, u_k^*).$$

The explicit expressions of u_k^* and w_k^* can be obtained by seeking fixed points of the above:

$$u_k^* = \varphi_{1,k}(x_k, \varphi_{2,k}(x_k, u_k^*)) = -P_{u_k} x_k$$
$$w_k^* = \varphi_{2,k}(x_k, \varphi_{1,k}(x_k, w_k^*)) = P_{w_k} x_k,$$

which is (3.9) and (3.10). Then the pair of (3.9) and (3.10) for each k constitutes a saddle point for the static zero-sum game at k, and by substituting (3.9) and (3.10) into (3.11) (or (3.12)), we arrive at the GRE. Proceeding similarly, we can obtain the state feedback saddle-point strategies in (ii) with the GRE for all k, where the corresponding saddle-point value is $V_0(x_0)$.

To prove part (iii), suppose that it has a negative eigenvalue for some $\bar{k} \in [0, N-1]$. Then the corresponding static zero-sum game does not admit a saddle point. In fact, there exists a sequence of maximizer strategies by which the upper value of this static zero-sum game becomes unbounded at $\bar{k}$, which also proves the necessity of part (i) [106]. This completes the proof. □

3.3.2 Infinite-Horizon Case

We now discuss the infinite-horizon problem of state feedback minimax control over the TCP-network. Before presenting the result, we provide some preliminaries. In this section, we assume that $Q_N = 0$. We first state the infinite-horizon version of the solution in Lemma 3.1.

- *The associated generalized algebraic Riccati equation (GARE) can be written as*

$$\bar{Z} = Q + \bar{P}_u^T(\alpha R + \bar{\alpha} B^T \bar{Z} B)\bar{P}_u - \gamma^2 \bar{P}_w^T \bar{P}_w + \bar{H}^T \bar{Z} \bar{H}, \tag{3.13}$$

  where $\bar{H}$, $\bar{P}_u$, and $\bar{P}_w$ are infinite-horizon versions of (3.8) with respect to $\bar{Z}$.

- *The stationary minimax controller and the worst-case disturbance are*

$$\bar{u}_k^* = -\bar{P}_u x_k \tag{3.14}$$
$$\bar{w}_k^* = \bar{P}_w x_k. \tag{3.15}$$

- *The infinite-horizon version of the existence condition in Lemma 3.1(i) is given by*

$$\rho(D^T \bar{Z} D) < \gamma^2. \tag{3.16}$$

We also introduce the time-reverse notation, which is used in the next proposition that states the convergence of the GRE. Let $\tilde{Z}_k := Z_{N-k}$. Then the GRE in (3.7) can be rewritten as

$$\tilde{Z}_{k+1} = Q + \tilde{P}_{u_k}^T(\alpha R + \bar{\alpha} B^T \tilde{Z}_k B)\tilde{P}_{u_k} - \gamma^2 \tilde{P}_{w_k}^T \tilde{P}_{w_k} + \tilde{H}_k^T \tilde{Z}_k \tilde{H}_k, \tag{3.17}$$

where $\tilde{P}_{u_k}$, $\tilde{P}_{w_k}$, and $\tilde{H}_k$ are the time-reverse versions of (3.8) in Lemma 3.1 with respect to $\tilde{Z}_k$. The time-reverse version of the concavity condition can be written as

$$\rho(D^T \tilde{Z}_k D) < \gamma^2. \tag{3.18}$$

Then we have the following result:

Proposition 3.1. *Suppose* (A, B) *is controllable and* $(A, Q^{1/2})$ *is observable. Define the sets*

$$\Gamma_1(\alpha) := \{\gamma > 0 : \bar{Z} \geq 0 \textit{ solves (3.13) and satisfies (3.16)}\}$$
$$\Lambda_1(\gamma) := \{\alpha \in [0, 1) : \bar{Z} \geq 0 \textit{ solves (3.13) and satisfies (3.16)}\}.$$

Let $\gamma_1^*(\alpha) := \inf\{\gamma : \ \gamma \in \Gamma_1(\alpha)\}$ *and* $\alpha_c(\gamma) := \inf\{\alpha : \ \alpha \in \Lambda_1(\gamma)\}$*. Then for any finite* $\gamma > \gamma_1^*(\alpha)$ *and* $\alpha > \alpha_c(\gamma)$*, as* $k \to \infty$*,* $\{\tilde{Z}_k\} \to \bar{Z}^+$ *where* $\bar{Z}^+$ *is a fixed point of (3.13) that satisfies (3.16).*

Proof. Let us first note some basic facts regarding the GARE in (3.13). In [106], it was proven that when $\alpha = 1$, (3.16) is a necessary and sufficient condition that guarantees convergence of the GRE in (3.17). In particular, for a fixed $\gamma > \gamma_1^*(1)$, given a fixed point of (3.13) that satisfies (3.16), $\{\tilde{Z}_k\}$ converges to $\bar{Z}^+$. Now, when $\alpha = 0$, (3.13) can be written as

$$\bar{Z} = A^T \bar{Z} A + Q + A^T \bar{Z} D(\gamma^2 I - D^T \bar{Z} D)^{-1} D^T \bar{Z} A, \tag{3.19}$$

which is the algebraic Riccati equation (ARE) associated with the optimization problem of (B.1) in Appendix B.1. If A is stable, (3.19) has a solution that satisfies (3.16), which is also equivalent to saying that $\{\tilde{Z}_k\}$ converges to $\bar{Z}^+$ [106]. When A is unstable, since the maximum cost of (B.1) is not bounded, (3.19) does not admit any solution in the class of positive semi-definite matrices, which also shows that $\{\tilde{Z}_k\}$ does not converge for any $\gamma > 0$. Thus, $\Gamma(0)$ is an empty set when A is unstable.

Now, from definitions of $\gamma_1^*(\alpha)$ and $\alpha_c(\gamma)$, $\bar{Z} \geq 0$ is a solution to (3.13) that satisfies (3.16). Due to Lemma B.1(ii), $\bar{Z}$ constitutes an upper bound on the GRE. Therefore, we have (3.18), which guarantees monotonicity of the GRE from Lemma B.1(i). Then, we can conclude that the monotonic and bounded sequence $\{\tilde{Z}_k\}$ converges as $k \to \infty$. This completes the proof. □

We also have the following result which shows the relationship between LQG and minimax control over the TCP-network.

Proposition 3.2. *Suppose that the assumptions in Proposition 3.1 hold. Then, as $\gamma \to \infty$, $\bar{Z}^+$ defined in Proposition 3.1 converges to the solution of the following ARE:*

$$\bar{Z}^+ = A^T \bar{Z}^+ A - \alpha A^T \bar{Z}^+ B(R + B^T \bar{Z}^+ B)^{-1} B^T \bar{Z}^+ A + Q.$$

Proof. The value of the soft-constrained zero-sum dynamic game decreases in γ [106]. Then the result follows immediately. □

We now state the main result of this section.

Theorem 3.1. *Suppose (A, B) is controllable and $(A, Q^{1/2})$ is observable. Then for any finite $\gamma > \gamma_1^*(\alpha)$ and $\alpha > \alpha_c(\gamma)$, the following hold:*

(i) The state feedback minimax controller is given by (3.14) with $\bar{Z}^+$.

(ii) Suppose $\alpha \bar{P}_u^T R \bar{P}_u - \gamma^2 \bar{P}_w^T \bar{P}_w > 0$. Then the closed-loop system with the worst-case disturbance in (3.15), i.e., $x_{k+1} = (A - \alpha_k B\bar{P}_u + D\bar{P}_w)x_k$, is stable in the mean-square sense, that is, $\mathbb{E}\{|x_k|^2\} \to 0$ as $k \to \infty$ for all initial conditions.

(iii) The closed-loop system, i.e., $x_{k+1} = (A - \alpha_k B\bar{P}_u)x_k + Dw_k$, is bounded in the mean-square sense, that is, there exists $M \geq 0$ such that $\mathbb{E}\{|x_k|^2\} \leq M$ for all k and initial conditions.

(iv) *The state feedback minimax controller achieves the performance level of γ, that is,* $\ll \mathcal{T}^{\infty}_{\mu^*_\gamma} \gg \leq \gamma$.

Proof. Parts (i) follows from Proposition 3.1. To prove part (ii), by using (3.13), we have

$$\begin{aligned}\mathbb{E}\{|x_{k+1}|^2_{\bar{Z}^+}\} - \mathbb{E}\{|x_k|^2_{\bar{Z}^+}\} &= \mathbb{E}\{x_k^T(\bar{H}^T\bar{Z}^+\bar{H} + \bar{\alpha}\bar{P}_u^T B^T\bar{Z}^+ B\bar{P}_u - \bar{Z}^+)x_k\} \\ &= -\mathbb{E}\{x_k^T Q x_k + \alpha x_k^T \bar{P}_u^T R\bar{P}_u x_k - \gamma^2 x_k^T \bar{P}_w^T \bar{P}_w x_k\}.\end{aligned}$$

Now, we have

$$\mathbb{E}\{x_{k+1}^T\bar{Z}^+x_{k+1}\} = x_0^T\bar{Z}^+x_0 - \sum_{i=0}^{k}\mathbb{E}\Big\{x_i^T(Q + \alpha\bar{P}_u^T R\bar{P}_u - \gamma^2\bar{P}_w^T\bar{P}_w)x_i\Big\}.$$

Since the left-hand side of the above equation is bounded below by zero, we have

$$\lim_{k\to\infty}\mathbb{E}\{x_k^T(Q + \alpha\bar{P}_u^T R\bar{P}_u - \gamma^2\bar{P}_w^T\bar{P}_w)x_k\} = 0.$$

Then in view of the observability assumption and $(\alpha\bar{P}_u^T R\bar{P}_u - \gamma^2\bar{P}_w^T\bar{P}_w) > 0$, we must have $\mathbb{E}\{|x_k|^2\} \to 0$ as $k \to \infty$.

For part (iii), when $w_k \equiv 0$, we have the following equation:

$$Z = \alpha\bar{P}_u^T R\bar{P}_u + (1-\alpha)A^T Z A + \alpha(A - B\bar{P}_u)^T Z(A - B\bar{P}_u) + Q,$$

where $Z \geq 0$ exists due to Theorem 3 in [19] and the relationship between the minimax control when $w_k \equiv 0$ and the LQG control. Then from (ii), we can show that $\mathbb{E}\{|x_k|^2\} \to 0$ as $k \to \infty$ for all initial conditions when $w_k \equiv 0$; hence, the result follows.

To prove part (iv), first note that for any finite $\gamma > \gamma_1^*(\alpha)$, since the upper value of the game is bounded with $\bar{Z}^+$, we have the following inequality for all disturbances in ℓ_2^p:

$$J^{\infty}(\mu^*_\gamma, w) \leq x_0^T\bar{Z}^+x_0 + \gamma^2\mathbb{E}\Big\{\sum_{k=0}^{\infty}|w_k|^2\Big\}.$$

By taking $x_0 = 0$, the result holds. This completes the proof. □

3.4 Minimax Estimation over the TCP-Network

This section considers minimax estimation for the TCP-case, which corresponds to part (b) of the certainty equivalence principle in Appendix B.2. The detailed analysis of this section can be found in Chapter 2.

Lemma 3.2. *Consider the zero-sum dynamic game in (3.5) with $\alpha \in [0,1]$, $\beta \in [0,1]$, and a fixed $\gamma > 0$. Then:*

(i) A stochastic minimax estimator (SME) exists if and only if

$$\rho(\Sigma_k Q) < \gamma^2, \ \forall k \in [0, N-1], \tag{3.20}$$

where Σ_k is generated by the following generalized stochastic Riccati equation (GSRE) in (2.14).

(ii) The SME is

$$\bar{x}_{k+1} = A\bar{x}_k + \alpha_k B u_k + A\Pi_k(\gamma^{-2} Q \bar{x}_k + \beta_k C^T V^{-1}(y_k - C\bar{x}_k)), \tag{3.21}$$

where the estimator gain $\Pi_k = (\Sigma_k^{-1} - \gamma^{-2} Q + \beta_k C^T V^{-1} C)^{-1}$.

We now construct the smallest values of γ and β for which the SME exists.

Proposition 3.3. *Suppose that (A, C) is observable and (A, D) is controllable. Define the following sets and parameters:*

$$\Gamma_2(\beta) := \{\gamma > 0 \ : \ \rho(\Sigma_k Q) < \gamma^2, \ \forall k\}, \quad \gamma_2^*(\beta) := \inf\{\gamma \ : \ \gamma \in \Gamma_2(\beta)\}$$
$$\Lambda_2(\gamma) := \{\beta \in [0,1) \ : \ \rho(\Sigma_k Q) < \gamma^2, \ \forall k\}, \quad \beta_c(\gamma) := \inf\{\beta \ : \ \beta \in \Lambda_2(\gamma)\}.$$

Then, for any finite $\gamma > \gamma_2^*(\beta)$ and $\beta > \beta_c(\gamma)$, $\rho(\Sigma_k Q) < \gamma^2$ holds for all k; hence, the SME exists.

Remark 3.1. *(i) Due to the acknowledgment nature of the TCP-case, the SME is a function of control and measurement packet loss information, i.e. $\{\alpha_k\}$ and $\{\beta_k\}$.*

(ii) The SME is time varying and random because the estimator gain depends on the GSRE that is a function of the measurement arrival process. Furthermore, when $\beta_k = 0$, while the Kalman filter in [21] is

identical to the open-loop estimator, the SME performs the state estimation under the worst-case disturbance.

(iii) For any finite $\gamma > \gamma_2^*(\beta)$ and $\beta > \beta_c(\gamma)$, by induction, we can show that $P_k \leq \Sigma_k$ for all k, where P_k with $P_0 = Q_0^{-1}$ is the error covariance matrix of the Kalman filter in Appendix A.3. Moreover, as $\gamma \to \infty$, the SME and Σ_k converge to the Kalman filter and P_k in [21].

3.5 Minimax Control over the TCP-Network

In this section, we consider part (c) of the certainty equivalence principle in Appendix B.2 and therefore complete the design of the output feedback minimax control system over the TCP-network. Toward this end, we combine the results in Sections 3.3 and 3.4, and then introduce one additional existence condition for the worst-case state estimator.

Theorem 3.2 ([99, 101])**.** *For any γ, α, and β, suppose (3.6) and (3.20) hold for all k, i.e. there exist the state feedback minimax controller and the SME over the TCP-network. Then:*

(i) The worst-case state estimator, $\hat{x}_k$, exists if

$$\rho(\Sigma_k Z_k) < \gamma^2, \text{ for all } k \in [0, N-1]. \tag{3.22}$$

(ii) If the condition in (i) holds, then the worst-case state estimator can be written as

$$\hat{x}_k = (I - \gamma^{-2}\Sigma_k Z_k)^{-1}\bar{x}_k, \tag{3.23}$$

where $\bar{x}_k$ is generated by the SME in Lemma 3.2.

(iii) If the condition in (i) holds, then the output feedback minimax controller is given by (3.9) with (3.23). Furthermore, this controller achieves the disturbance attenuation performance level of γ, that is, $\ll \mathcal{T}_{\mu_\gamma^*}^N \gg \leq \gamma$.

Remark 3.2. *As expected from standard H^∞ control theory, there are three conditions on γ; (3.20) is for the existence of the SME, (3.6) is related to the state feedback minimax controller, and (3.22) is the spectral radius condition* *that ensures the existence of the worst-case state estimator. Moreover, unlike the LQG case considered in [21], there is no separation between control and estimation due to (3.22).*

For the infinite-horizon case, we can use the theories developed in Sections 3.3 and 3.4 to obtain a corresponding output feedback minimax controller. This is done next; the proof is similar to that of Theorem 3.2.

Theorem 3.3. *Suppose that (A, B) and (A, D) are controllable, and $(A, Q^{1/2})$ and (A, C) are observable. Define*

$$\Gamma_3(\alpha, \beta) := \{\gamma > 0 \ : \ \gamma > \gamma_1^*(\alpha), \ \gamma > \gamma_2^*(\beta), \ \rho(\Sigma_k \bar{Z}^+) < \gamma^2, \forall k\},$$

where $\bar{Z}^+$ is the solution of the GARE in (3.13) that satisfies (3.16). Let $\gamma_3^*(\alpha, \beta) := \inf\{\gamma \ : \ \gamma \in \Gamma_3(\alpha, \beta)\}$. Then for any finite $\gamma > \gamma_3^*(\alpha, \beta)$, $\alpha > \alpha_c(\gamma)$, and $\beta > \beta_c(\gamma)$, the stationary output feedback minimax controller is given by (3.14) with the following worst-case state estimator:

$$\hat{x}_k = (I - \gamma^{-2}\Sigma_k \bar{Z}^+)^{-1} \bar{x}_k, \tag{3.24}$$

where $\bar{x}_k$ is generated by the SME in Lemma 3.2. Finally, $\ll \mathcal{T}_{\mu_\gamma^*}^{\infty} \gg \leq \gamma$. □

Remark 3.3. *(i) $\gamma_3^*(\alpha, \beta)$ is the smallest value of γ that satisfies all the existence conditions, which is the optimum disturbance attenuation level of the original disturbance attenuation problem.*

(ii) The optimum disturbance attenuation level is a function of α and β. In fact, $\gamma_3^*(1, 1)$ is related to the deterministic H^∞ control problem, and $\gamma_3^*(0, 0)$ is analogous to the open-loop problem that is not finite when A is unstable.

(iii) As can be seen from Theorem 3.3, the critical values, $\alpha_c(\gamma)$ and $\beta_c(\gamma)$, are coupled with each other through γ; hence, in general, their values are problem dependent and cannot be quantified analytically. This fact actually stems from standard H^∞ control, in which the optimum disturbance attenuation level (the smallest value of γ in the context of standard H^∞ control) cannot be determined analytically, and a heuristic approach is generally used depending on the problem at hand [106, 111].

(iv) *When $\gamma \to \infty$, from (3.24), we can easily see that $\hat{x}_k = \bar{x}_k$ for all k. Furthermore, the state feedback minimax controller as well as the SME collapse to the LQG system presented in [21] in view of Proposition 3.2 and Remark 3.1(iii).*

3.6 Minimax Control over the UDP-Network

In this section, we study the minimax control problem over the UDP-network.

3.6.1 Finite-horizon Case

We first consider the finite-horizon problem. The UDP-like information structure and the associated cost function are given by (3.3) and (3.5), respectively. We assume that the linear dynamical system in (3.1) has no measurement noise ($E = 0$), and C is the identity matrix, that is, in case of transmission the controller has perfect access to instantaneous value of the state. We then have the following linear dynamical system:

$$x_{k+1} = Ax_k + \alpha_k Bu_k + Dw_k,\ y_k = \beta_k x_k. \tag{3.25}$$

We let $\bar{\alpha} := \alpha - \alpha^2$, $\alpha' := 1 - \alpha$ and $\beta' := 1 - \beta$. We now obtain the output feedback minimax controller (and the worst-case disturbance) under the UDP-type information structure for (3.25).

Lemma 3.3. *Consider the zero-sum dynamic game in (3.5) with (3.25). For fixed $\gamma > 0$, $\alpha \in [0,1]$ and $\beta \in [0,1]$, we have the following result:*

(i) *A unique output feedback saddle-point solution exists if and only if*

$$\rho(D^T Z_{k+1} D) < \gamma^2, \tag{3.26}$$

where Z_k is generated by the following coupled generalized Riccati equa-

tions (GREs): $Z_N = Q_N$, $U_N = 0$ and

$$\begin{aligned}
Z_k &= \breve{H}_k^T Z_{k+1} \breve{H}_k + Q - \gamma^2 \breve{P}_{w_k}^T \breve{P}_{w_k} \\
&\quad + \breve{P}_{u_k}^T (\alpha R + \bar{\alpha} B^T Z_{k+1} B + \beta' \bar{\alpha} B^T U_{k+1} B) \breve{P}_{u_k} \\
&= Q + A^T Z_{k+1} A - U_k + \beta' A^T U_{k+1} A \qquad (3.27) \\
U_k &= \beta' A^T U_{k+1} A + \breve{P}_{w_k}^T (\gamma^2 I - D^T Z_{k+1} D) \breve{P}_{w_k} + 2\alpha \breve{P}_{u_k}^T B^T Z_{k+1} D \breve{P}_{w_k} \\
&\quad - \breve{P}_{u_k}^T (\alpha R + \beta' \bar{\alpha} B^T U_{k+1} B) \breve{P}_{u_k} + 2\alpha \breve{P}_{u_k}^T B^T Z_{k+1} A - 2 \breve{P}_{w_k}^T B^T Z_{k+1} A, \qquad (3.28)
\end{aligned}$$

where

$$\begin{aligned}
\breve{H}_k &= A - \alpha B \breve{P}_{u_k} + D \breve{P}_{w_k} \\
\breve{P}_{u_k} &= (S_k + \alpha B^T Z_{k+1} D M_k^{-1} D^T Z_{k+1} B)^{-1} \\
&\quad \times B^T (I + Z_{k+1} D M_k^{-1} D^T) Z_{k+1} A \\
\breve{P}_{w_k} &= (M_k + \alpha D^T Z_{k+1} B S_k^{-1} B^T Z_{k+1} D)^{-1} \\
&\quad \times D^T (I - \alpha Z_{k+1} B S_k^{-1} B^T) Z_{k+1} A \\
S_k &= R + B^T (Z_{k+1} + \alpha' \beta' U_{k+1}) \\
M_k &= \gamma^2 I - D^T Z_{k+1} D.
\end{aligned}$$

(ii) The saddle-point solution can be written as, where $\hat{x}_k = \mathbb{E}\{x_k | \mathcal{G}_k\}$:

$$u_k^* = -\breve{P}_{u_k} \hat{x}_k \qquad (3.29)$$

$$w_k^* = \breve{P}_{w_k} \hat{x}_k, \;\; k \in [0, N-1]. \qquad (3.30)$$

Proof. To prove parts (i) and (ii), we need to employ dynamic programming or rather the Isaacs equation. At stage N, the value function is given by $V_N(x_N) = \mathbb{E}\{x_N^T Q_N x_N | \mathcal{G}_N\}$. It is easy to see that, from the dynamic programming equation, the cost-to-go from stage $N-1$ can be expressed as

$$\begin{aligned}
V_{N-1}(x_{N-1}) &= \min_{u_{N-1}} \max_{w_{N-1}} \mathbb{E}\Big\{ h_{N-1}(x,u,w) + V_N(x_N) | \mathcal{G}_{N-1} \Big\} \qquad (3.31) \\
&= \max_{w_{N-1}} \min_{u_{N-1}} \mathbb{E}\Big\{ h_{N-1}(x,u,w) + V_N(x_N) | \mathcal{G}_{N-1} \Big\} \qquad (3.32) \\
&= \mathbb{E}\{ |x_{N-1}|_{Z_{N-1}}^2 + |e_{N-1}|_{U_{N-1}}^2 | \mathcal{G}_{N-1} \}, \qquad (3.33)
\end{aligned}$$

where $h_{N-1}(x,u,w) := |x_{N-1}|_Q^2 + \alpha_{N-1} |u_{N-1}|_R^2 - \gamma^2 |w_{N-1}|^2$, Z_{N-1} is the GRE in Lemma 3.1, $e_k := x_k - \hat{x}_k$ with $\hat{x}_k = \mathbb{E}\{x_k|\mathcal{G}_k\}$, and U_{N-1} is given by

$$\begin{aligned}U_{N-1} = &-\alpha P_{u_{N-1}}^T(R+B^TQ_NB)P_{u_{N-1}} + P_{w_{N-1}}^T(\gamma^2I - D^TQ_ND)P_{w_{N-1}} \\ &+ 2\alpha P_{u_{N-1}}^T B^TQ_NA + 2\alpha P_{u_{N-1}}^T B^TQ_NDP_{w_{N-1}} - 2P_{w_{N-1}}D^TQ_NA,\end{aligned}$$

where $P_{u_{N-1}}$ and $P_{w_{N-1}}$ are defined in Lemma 3.1. Note that $U_{N-1} \geq 0$. The equality in (3.33) is achieved by using the following saddle-point solution under (3.26):

$$u_{N-1}^* = -P_{u_{N-1}}\hat{x}_{N-1},\ w_{N-1}^* = P_{w_{N-1}}\hat{x}_{N-1},$$

which can be achieved by solving the static zero-sum game in (3.31).

Note that as mentioned in [19], the estimator error becomes a function of u_{N-2} at stage $N-2$ so that there is dual effect. To see this, we first write the cost-to-go from stage $N-2$:

$$V_{N-2}(x_{N-2}) = \min_{u_{N-2}}\max_{w_{N-2}}\mathbb{E}\big\{h_{N-2}(x,u,w) + V_{N-1}(x_{N-1})|\mathcal{G}_{N-2}\big\} \tag{3.34}$$

$$= \max_{w_{N-2}}\min_{u_{N-2}}\mathbb{E}\big\{h_{N-2}(x,u,w) + V_{N-1}(x_{N-1})|\mathcal{G}_{N-2}\big\}. \tag{3.35}$$

Now, note that the estimation error at $k = N-2$ is zero when $\beta_{N-2} = 1$, and $e_{N-1} = Ae_{N-2} + (\alpha_{N-2} - \alpha)u_{N-2}$, otherwise. Therefore, (3.34) (or (3.35)) yields a unique minimizer and maximizer under (3.26), which can be written as follows:

$$\begin{aligned}u_{N-2}^* &= -(R + B^T(Z_{N-1} + \alpha'\beta'U_{N-1})B)^{-1}B^TZ_{N-1}(A\hat{x}_{N-2} + Dw_{N-2}^*) \\ &=: \psi_{1,N-2}(\hat{x}_{N-2}, w_{N-2}^*) \\ w_{N-2}^* &= (\gamma^2I - D^TZ_{N-1}D)^{-1}D^TZ_{N-1}(A\hat{x}_{N-2} + \alpha Bu_{N-2}^*) \\ &=: \psi_{2,N-2}(\hat{x}_{N-2}, u_{N-2}^*).\end{aligned}$$

We obtain the saddle point, (u_{N-2}^*, w_{N-2}^*), for (3.34) (or (3.35)) by solving the above fixed-point equations:

$$u_{N-2}^* = \psi_{1,N-2}(\hat{x}_{N-2}, \psi_{2,N-2}(\hat{x}_{N-2}, u_{N-2}^*)) = -\breve{P}_{u_{N-2}}\hat{x}_{N-2} \tag{3.36}$$

$$w_{N-2}^* = \psi_{2,N-2}(\hat{x}_{N-2}, \psi_{1,N-2}(\hat{x}_{N-2}, w_{N-2}^*)) = \breve{P}_{w_{N-2}}\hat{x}_{N-2}. \tag{3.37}$$

Substituting (3.36) and (3.37) into (3.34) (or (3.35)), we obtain

$$V_{N-2}(x_{N-2}) = \mathbb{E}\{|x_{N-2}|^2_{Z_{N-2}}|\mathcal{G}_{N-2}\} + \mathbb{E}\{|e_{N-2}|^2_{U_{N-2}}|\mathcal{G}_{N-2}\},$$

where Z_{N-2} and U_{N-2} are given in (3.27) and (3.28), respectively. Then proceeding similarly, the minimax controller and the worst-case disturbance that constitute a saddle point can be written as (3.29) and (3.30), respectively. This completes the proof. □

In summary, for the linear system given in (3.25), the corresponding minimax controller is (3.29), which is linear in the information $\mathcal{G}_k$ given by (3.3), and is a function of two coupled nonlinear GREs (3.27) and (3.28). If the above existence condition fails to hold, then the minimax controller does not exist. In fact, the value of the corresponding zero-sum dynamic game would then be infinite as we discussed in the TCP-case.

It should be mentioned that there is no known general solution to the problem of LQG control over the UDP-network for the noisy measurement model in (3.1b), since the associated optimization problem is then no longer convex, and the optimal LQG controller is generally nonlinear in the available information ([21]). Also, there is no separation between control and estimation. We would naturally expect a similar difficulty to arise in the minimax control problem under the noisy measurement case.

We next proceed with the infinite-horizon case for again the additive noise free problem.

3.6.2 Infinite-Horizon Case

The infinite-horizon versions of the coupled GREs and the existence condition are provided below:

- *The coupled GAREs are given by*

$$\begin{aligned} Z &= \breve{H}^T Z \breve{H} + Q - \gamma^2 \breve{P}_w^T \breve{P}_w + \breve{P}_u^T(\alpha R + \bar{\alpha} B^T Z B + \beta' \bar{\alpha} B^T U B)\breve{P}_u \\ &= Q + A^T Z A - U + \beta' A^T U A \qquad (3.38) \\ U &= \beta' A^T U A + \breve{P}_w^T(\gamma^2 I - D^T Z D)\breve{P}_w - \breve{P}_u^T(\alpha R + \beta' \bar{\alpha} B^T U B)\breve{P}_u \\ &\quad + 2\alpha \breve{P}_u^T B^T Z A - 2\breve{P}_w^T B^T Z A + 2\alpha \breve{P}_u^T B^T Z D \breve{P}_w, \qquad (3.39) \end{aligned}$$

where P'_u, P'_w, and H' are infinite-horizon versions of P'_{u_k}, P'_{w_k}, and H'_k, respectively.

- *The corresponding minimax controller and the worst-case disturbance are given by*

$$u_k^* = -\breve{P}_u \hat{x}_k \quad (3.40)$$

$$w_k^* = -\breve{P}_w \hat{x}_k. \quad (3.41)$$

- *The existence condition can be written as*

$$\rho(D^T Z D) < \gamma^2. \quad (3.42)$$

We need to obtain conditions on γ, α and β that guarantee convergence of the coupled GREs in (3.27) and (3.28) under (3.26), which can be characterized by

$$\begin{aligned}
\gamma_U^*(\alpha,\beta) = \inf\{\gamma > 0 : \lim_{k\to\infty} \tilde{Z}_k = Z, \lim_{k\to\infty} \tilde{U}_k = U,\ Z \geq 0 \text{ and } U \geq 0 \text{ solve}\\
\text{(3.38) and (3.39), and satisfy (3.42)}\}\\
\alpha_c^U(\gamma,\beta) = \inf\{\alpha \in [0,1) : \lim_{k\to\infty} \tilde{Z}_k = Z, \lim_{k\to\infty} \tilde{U}_k = U,\ Z \geq 0 \text{ and } U \geq 0\\
\text{solve (3.38) and (3.39), and satisfy (3.42)}\}\\
\beta_c^U(\gamma,\alpha) = \inf\{\beta \in [0,1) : \lim_{k\to\infty} \tilde{Z}_k = Z, \lim_{k\to\infty} \tilde{U}_k = U,\ Z \geq 0 \text{ and } U \geq 0\\
\text{solve (3.38) and (3.39), and satisfy (3.42)}\},
\end{aligned}$$

where $\tilde{Z}_k$ and $\tilde{U}_k$ are the time-reverse equations of (3.27) and (3.28), respectively, as introduced in Section 3.3.2. Note that these parameters are coupled with each other. Also, if $\gamma > \gamma_U^*(\alpha,\beta)$, $\alpha > \alpha_c^U(\gamma,\beta)$, and $\beta > \beta_c^U(\gamma,\alpha)$, then the infinite-horizon minimax controller for the UDP-case is (3.40), provided that γ is finite, which stabilizes the closed-loop system and achieves the disturbance attenuation level of γ. Moreover, as $\gamma \to \infty$, the critical values, $\alpha_c^U(\gamma,\beta)$ and $\beta_c^U(\gamma,\alpha)$, converge to the corresponding LQG values in [19, 21].

For the LQG problem, the explicit convergence conditions of the corresponding Riccati equations were obtained in [19, 21] when B is invertible. Since the minimax controller is equivalent to the LQG controller when γ asymptotically goes to infinity, those conditions are necessary for the minimax controller; that is, due to the existence condition, the conditions in [19, 21] are only necessary for the convergence of (3.27) and (3.28) even if B is invertible. We should note that the general convergence conditions cannot be obtained analytically, because the critical values, $\alpha_c^U(\gamma, \beta)$ and $\beta_c^U(\gamma, \alpha)$, are coupled with each other.

We now state the main result of this section.

Theorem 3.4. *Suppose that (A, B) and (A, D) are controllable, and $(A, Q^{1/2})$ is observable. Suppose that $\gamma > \gamma_U^*(\alpha, \beta)$ is finite, and $\alpha > \alpha_c^U(\gamma, \beta)$ and $\beta > \beta_c^U(\gamma, \alpha)$. Then:*

(i) The minimax controller is given by (3.40).

(ii) Suppose $\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w > 0$. Then the closed-loop system with the worst-case disturbance in (3.41) and the estimation error are bounded in the mean-square sense, that is, there exist $M, M' \geq 0$ such that $\mathbb{E}\{|x_k|^2\} \leq M$ and $\mathbb{E}\{|e_k|^2\} \leq M'$ for all k and initial conditions.

(iii) The closed-loop system with an arbitrary disturbance and the estimation error are bounded in the mean-square sense.

(iv) The minimax controller in (i) achieves the disturbance attenuation level of γ.

Proof. Parts (i) and (iv) follow from the preceding discussion. To prove part (ii), by using (3.38) and (3.39), we have

$$\begin{aligned}&\mathbb{E}\{|x_{k+1}|_Z^2 - |x_k|_Z^2 + |e_{k+1}|_U^2 - |e_k|_U^2\}\\&= \mathbb{E}\{-x_k^T(Q + \alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)x_k + e_k^T(\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)e_k\}.\end{aligned}$$

Summing up the above expression over k yields

$$\begin{aligned}&\mathbb{E}\{|x_{k+1}|_Z^2 + |e_{k+1}|_U^2\}\\&= \mathbb{E}\{|x_0|_Z^2 + |e_0|_U^2\} + \sum_{i=0}^{k} \mathbb{E}\{e_i^T(\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)e_i\}\\&\quad - \sum_{i=0}^{k} \mathbb{E}\{x_i^T(Q + \alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)x_i\}.\end{aligned}$$

Note that for any $L \geq 0$, $\beta'\mathbb{E}\{x_k^T L x_k\} \geq \mathbb{E}\{e_k^T L e_k\}$, and we have $\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w > 0$. Therefore,

$$\begin{aligned}\mathbb{E}\{|x_{k+1}|_Z^2 + |e_{k+1}|_U^2\} &\leq \mathbb{E}\{|x_0|_Z^2 + |e_0|_U^2\} \\ &\quad - \sum_{i=0}^{k} \mathbb{E}\{x_i^T (Q + \beta(\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)) x_i\}.\end{aligned}$$

Since the left-hand side of the above inequality is bounded below by zero, we have $\lim_{k\to\infty} \mathbb{E}\{x_k^T (Q + \beta(\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w)) x_k\} = 0$. Since $\alpha \breve{P}_u^T R \breve{P}_u - \gamma^2 \breve{P}_w^T \breve{P}_w > 0$, in view of the observability assumption, we have the desired result.

For part (iii), when $w_k \equiv 0$, we have

$$\begin{aligned}Z &= \alpha' A^T Z A + \alpha(A - B\breve{P}_u)^T Z (A - B\breve{P}_u) + \breve{P}_u^T (\alpha R + \bar{\alpha}\beta' B^T U B)\breve{P}_u + Q \\ U &= \alpha A^T Z A - \alpha(A - B\breve{P}_u)^T Z (A - B\breve{P}_u) + \beta' A^T U A \\ &\quad - \breve{P}_u^T (\alpha R + \bar{\alpha}\beta' B^T U B)\breve{P}_u,\end{aligned}$$

where $Z \geq 0$ and $U \geq 0$ exist due to Theorem 9 in [19] and the relationship between the minimax control when $w_k \equiv 0$ and the LQG control. Then from (ii), we can show that $\mathbb{E}\{|x_k|^2\}$ and $\mathbb{E}\{|e_k|^2\}$ are bounded when $w_k \equiv 0$; hence, the result follows. □

3.7 Numerical Examples

In this section, we provide numerical examples to demonstrate the relationship between α, β, and γ, and compare the disturbance attenuation performance for different values of γ.

3.7.1 Stability and Performance Region

Consider the following system:

$$x_{k+1} = A x_k + \alpha_k u_k + w_k, \tag{3.43}$$

where $A = 2$ and $A = 1.1$. We take $R = 1$ and $Q = 1$.

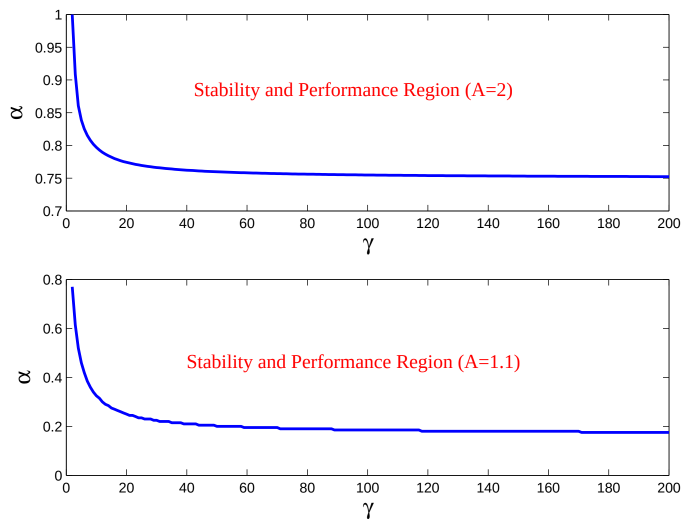


Figure 3.1: Stability and performance region of (3.43) for the TCP-case.

Figure 3.1 shows the stability and performance region of (3.43) for the TCP-case. To obtain the region numerically, we use the following approach:

(S.1) Fix $\alpha = 1$ and take a sufficiently large value of $\gamma > 0$.

(S.2) Obtain the solution of the GARE in (3.13), and check the existence condition in (3.16).

(S.3) If the existence condition holds, decrease α and then go to (S.2). Otherwise, it is the critical value of α for that γ; go to the next step.

(S.4) Decrease γ and fix $\alpha = 1$. Go to (S.2).

Figure 3.1 shows (as goes with intuition) that the system needs a more reliable communication channel if the high level of disturbance attenuation is required. For both cases, $\alpha_c(\gamma) \to \alpha_R$ as $\gamma \to \infty$ where $\alpha_R = 1 - (1/2^2) = 0.75$ when $A = 2$ has been calculated in [19]. Moreover, as the plant becomes more open-loop unstable, the stability and performance region becomes smaller. This is an expected result, since the more open-loop unstable a plant is, the more frequently we need to measure its state and control it. Finally, it is easy to see that $\alpha > \alpha_R$ is a necessary condition for the existence of the state feedback minimax controller for the TCP-case.

The region of stability and performance for the UDP-case is shown in Fig. 3.2. We used the same approach above to obtain this plot. As expected, the

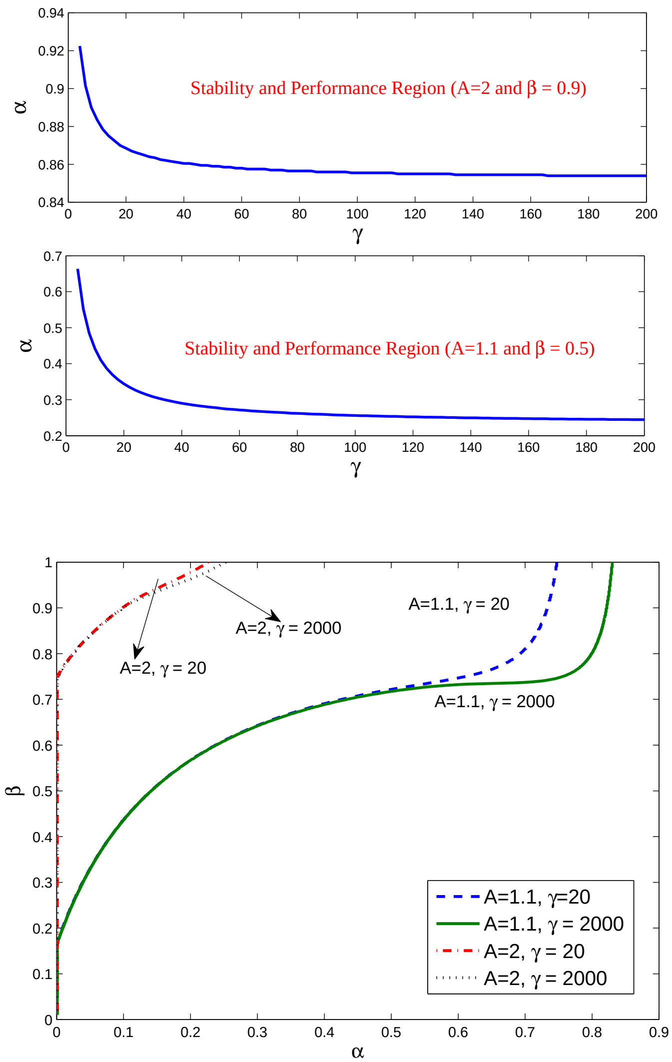


Figure 3.2: Stability and performance region of (3.43) for the UDP-case.

UDP controller has a smaller stability and performance region in terms of α, β, and γ than the TCP-case. Moreover, as $\gamma \to \infty$, $\alpha_c^U(\gamma, \beta)$ and $\beta_c^U(\gamma, \alpha)$ converge to the corresponding critical values shown in [19]. This also shows that the condition given in [19] is a necessary condition for the existence of the minimax controller for the UDP-case.

3.7.2 Disturbance Attenuation Performance (TCP-case)

We use the pendubot system as in [21], where the system and cost matrices can be found. Figure 3.3 shows the existence regions of the state feedback

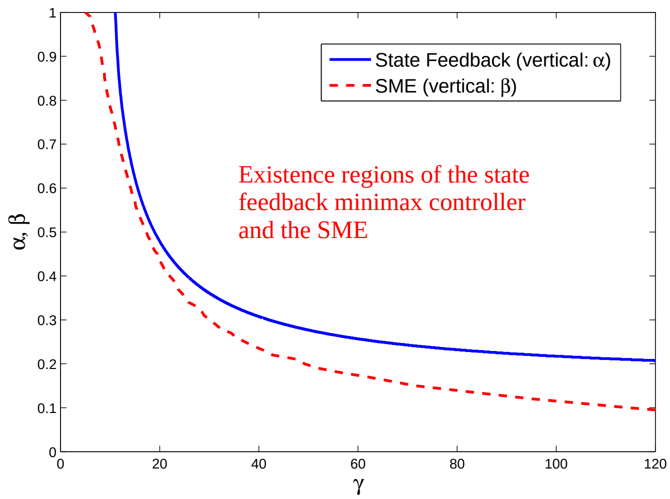


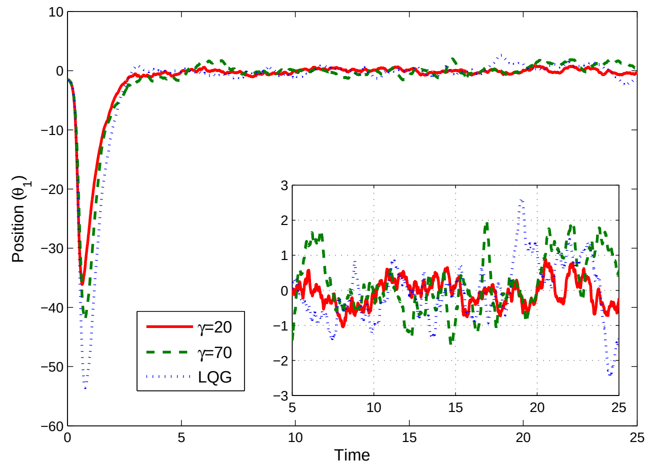
Figure 3.3: Existence regions of the state feedback minimax controller and the SME for the pendubot system.

minimax controller and the SME. This plot is also obtained by using the approach described in Section 3.7.1. The vertical axis is α for the state feedback controller, whereas it is β for the SME. Note that the intersection of regions above the two lines guarantees the existence of the state feedback minimax controller as well as the SME. Moreover, as $\gamma \to \infty$, all these regions converge to the corresponding value of the LQG problem in [21].

Figure 3.4 shows the disturbance attenuation performance of the minimax controller for different values of γ when w_k is Gaussian or a sinusoidal disturbance with amplitude of 0.01. We use $\alpha = 0.8$ and $\beta = 0.9$. As can be seen, when $\gamma = 20$, the minimax controller outperforms the LQG controller. Finally, when γ is sufficiently large, the performance of the minimax controller is identical to the corresponding LQG controller in [21].

3.8 Conclusions

In this chapter, we have studied the minimax control problem for LTI systems over unreliable communication channels. We have considered two different scenarios for the communication channels: the TCP-case and the UDP-case. Unlike the previous work, we have considered the situation when the sensor noise and the disturbance are not necessarily stochastic processes, but



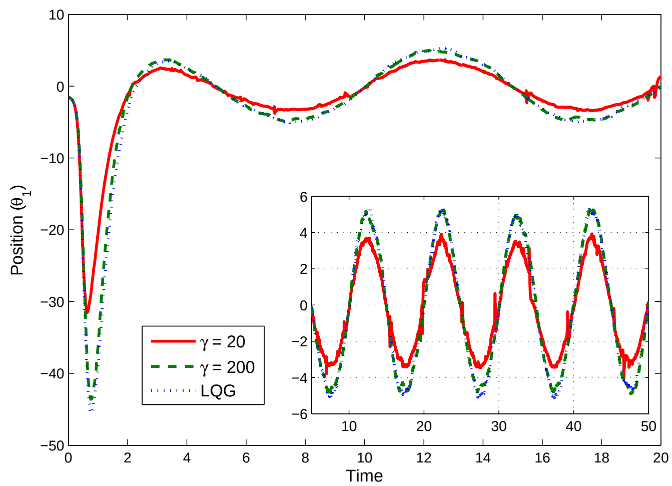


Figure 3.4: Disturbance attenuation performance with $\alpha = 0.8$ and $\beta = 0.9$ (top: Gaussian disturbance, bottom: sinusoidal disturbance).

are treated as adversarial inputs. The control problems are then naturally formulated within the framework of stochastic zero-sum dynamic games.

For both the TCP and UDP cases, we have obtained different classes of output feedback minimax controllers by characterizing the corresponding sets of existence conditions in terms of the H^∞ disturbance attenuation parameter and the packet loss rates. We have shown that stability and performance of the closed-loop system are determined by the disturbance attenuation parameter and the packet loss rates. Finally, as the disturbance attenuation parameter goes to infinity, the minimax controllers become equivalent to the corresponding LQG controllers (TCP or UDP controllers).

Part II

Mean Field Games

CHAPTER 4

RISK-SENSITIVE AND ROBUST MEAN FIELD GAMES

4.1 Introduction

In this chapter, we study two classes of mean field games; linear-quadratic risk-sensitive mean field game (LQ-RSMFG, **P1**) and LQ robust mean field game (LQ-RMFG, **P2**).

In **P1**, each agent minimizes an exponentiated performance index, which captures risk-sensitive behavior. In **P2**, each agent minimizes a worst-case performance index, where the performance is affected by a fictitious agent or an adversary who affects each agent's state system through an input (possibly independent across different agents' systems). In contrast to [113], the adversary in **P2** is state dependent, which is generally unknown to each agent. In both cases, that is **P1** and **P2**, we consider the heterogeneous agent case with infinite-horizon performance indices, and the individual agents are coupled with each other through the mean field term included in the individual performance indices.

Now, the main objectives of the chapter are as follows:

(i) to characterize Nash equilibria for **P1** and **P2** in a decentralized manner;

(ii) to establish conditions under which the Nash equilibria of **P1** and **P2** are *equivalent*;

(iii) to analyze limiting behaviors of equilibria in the large population regime as well as at particular limiting values of the design parameters.

To attain our goal, we use mean field game theory. Specifically, we first obtain an individual optimal decentralized controller by solving the LQ risk-sensitive optimal control problem (LQ-RSOCP) for **P1** or the LQ stochastic zero-sum differential game (LQ-SZSDG) for **P2**, which is a function of the local state information and an arbitrary deterministic function. Note that as shown in [75, 106, 78, 79, 77], LQ-RSOCP and LQ-SZSDG are equivalent; therefore, the corresponding optimal controllers are identical. We then construct a mean field system for **P1** and **P2** that is used to approximate the mass behavior effect on the individual agent by characterizing a unique deterministic function that can be determined off-line. This characterization is based on a fixed point analysis using a contraction mapping argument. Since the worst-case disturbance plays a crucial role in **P2**, these two mean field systems are not generally identical, which results in providing a different estimate of the mass behavior.

We show consistency between the best approximated mass behaviors for **P1** and **P2**, and the actual mass behavior in different senses. In particular, we prove that in the large population regime, the approximation error converges to zero in both the mean-square sense and the almost sure sense. We prove that the set of N-optimal decentralized controllers form ϵ-Nash equilibria for **P1** and **P2**, and ϵ can be made arbitrarily close to zero when the population size grows to infinity. We also show that the ϵ-Nash equilibria for **P1** and **P2** are *partially equivalent*, because of the equivalence of the individual optimal control laws, and the differences in the approximated mass behaviors obtained by the corresponding mean field systems.

Finally, we discuss limiting behaviors of the ϵ-Nash equilibria for **P1** and **P2** with respect to the design parameters. Specifically, we show that when the disturbance attenuation parameter goes to infinity, their ϵ-Nash equilibria are equivalent to that of the risk-neutral game considered in [71]. We also show by numerical examples that the Nash equilibria feature robustness. This follows from the inherent robustness property of the individual optimal control problems, i.e., LQ-RSOCP and LQ-SZSDG.

Organization

The chapter is organized as follows. In Section 4.2, we formulate LQ risk-sensitive mean field games and LQ robust mean field games. We solve the LQ risk-sensitive mean field game in Section 4.3, and discuss its limiting behaviors with respect to two design parameters in Section 4.4. In Section

4.5, we solve the LQ robust mean field game. In Section 4.6, the partial equivalence between the LQ risk-sensitive mean field game and the LQ robust mean field game, and their limiting behaviors, are discussed. In Section 4.7, numerical examples are presented to illustrate the results. We end the chapter with the concluding remarks of Section 4.8.

We use the following notation: universal constants c, M, ϵ, etc., are independent of the population size N and/or the time-horizon T. On the other hand, constants $\epsilon(N)$ and $X(T)$ are dependent on N and T, respectively, and $X(N,T)$ is dependent on N and T, jointly.

4.2 Problem Formulation

In this section, we formulate two problems: the linear-quadratic risk-sensitive mean field game (LQ-RSMFG, **P1**) and the linear-quadratic robust mean field game (LQ-RMFG, **P2**).

4.2.1 LQ Risk-Sensitive Mean Field Games (P1)

The stochastic differential equation (SDE) for agent i, $1 \leq i \leq N$, is given by

$$dx_i = (A(\theta_i)x_i + B(\theta_i)u_i)dt + \sqrt{\mu}D(\theta_i)dW_i(t), \tag{4.1}$$

where $x_i(0) = \bar{x}_i$; $x_i \in \mathbb{R}^n$ is the state; $u_i \in \mathbb{R}^m$ is the control input; $\{W_i(t), t \geq 0\}$ is a p-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$; and $\mu > 0$ is a noise intensity parameter. Further, $\theta_i \in \Theta \subset \mathbb{R}^q$, with $q := n(n+m+p)$, is an independent system vector that determines the triplet $(A(\theta), B(\theta), D(\theta))$ of each agent. We have the following assumption:

Assumption 4.1. *(a) $\{W_i(t), t \geq 0; 1 \leq i \leq N\}$ are independent across different agents.*

(b) $A(\cdot)$, $B(\cdot)$, and $D(\cdot)$ are continuous matrix-valued functions of θ with appropriate dimension for all $\theta \in \Theta$. Also, Θ is compact.

(c) $\bar{x}_i \in X$ for all i where X is a compact subset of $\mathbb{R}^n$.

(d) *For the first $N \geq 1$ agents, we have the following empirical distribution:*

$$F_N(\theta, x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{\theta_i \leq \theta,\ \bar{x}_i \leq x\}},\ \theta \in \Theta,\ x \in X,$$

where each inequality is componentwise.

(e) *There is a probability distribution $F(\theta, x)$ on $\mathbb{R}^{q+n}$ such that $F_N(\theta, x)$ converges weakly to $F(\theta, x)$ as $N \to \infty$. In other words, for any bounded and continuous function $p(\theta, x)$ on $\mathbb{R}^{q+n}$ [114],*

$$\lim_{N \to \infty} \int_{\mathbb{R}^{q+n}} p(\theta, x) dF_N(\theta, x) = \int_{\mathbb{R}^{q+n}} p(\theta, x) dF(\theta, x).$$

Remark 4.1. *(i) If each agent is uniform ($\theta_i = (A, B, D)$ and $\bar{x}_i = \bar{x}$ for all i), then $F(\theta, x)$ degenerates to a point mass, and in that case, (b)-(e) in Assumption 4.1 would not be needed.*

(ii) Let $F(\theta)$ and $F_N(\theta)$ be the marginal distribution functions of $F(\theta, x)$ and $F_N(\theta, x)$, respectively, with respect to θ. Due to the Glivenko-Cantelli theorem [108], $F_N(\theta)$ converges weakly when $\{\theta_i\}$ is a sequence of i.i.d. random variables under the distribution $F(\theta)$.

In this chapter, we consider only the risk averse problem; hence, $\delta > 0$. We let $\gamma := \sqrt{\delta / 2\mu}$ and call it the disturbance attenuation parameter. The significance and relevance of the disturbance attenuation parameter will be clarified throughout the chapter.

Now, given (4.1) and Assumption 4.1, in the problem under consideration, each agents seeks a controller that minimizes the following risk-sensitive performance index:

$$\mathbf{P1}:\ J_{1,i}^N(u_i, u_{-i}) = \limsup_{T \to \infty} \frac{\delta}{T} \log \mathbb{E}\Big\{ e^{\frac{1}{\delta}\phi_i^1(x, f_N, u)} \Big\}, \tag{4.2}$$

where $u_{-i} = \{u_1, ..., u_{i-1}, u_{i+1}, ..., u_N\}$, δ is the risk sensitivity index, and

$$\phi_i^1(x, f_N, u) := \int_0^T \|x_i(t) - f_N(t)\|_Q^2 + \|u_i(t)\|_R^2 dt, \tag{4.3}$$

where $Q \geq 0$ and $R > 0$. In (4.3), $f_N(t)$ denotes the mean field term or the mass behavior term that captures the average behavior of the first N agents:

$$f_N(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(t). \tag{4.4}$$

In view of this setting, the agents interact with each other through the mean field term, and the coupling effect on each agent is taken into account in (4.2).

Remark 4.2. *Let $\lambda = 1/\delta$. By using the Taylor expansion of (4.2) around $\lambda = 0$, one arrives at the following relation:*

$$J_{1,i}^N(u_i, u_{-i}) = \limsup_{T\to\infty} \frac{1}{T}\Big[\mathbb{E}\{\phi_i^1(x, f_N, u)\} + \frac{\lambda}{2} var\{\phi_i^1(x, f_N, u)\} + o(\lambda)\Big],$$

where var is the variance of $\phi_i^1(x, f_N, u)$ and $o(\lambda)$ is a higher order term (in λ) which involves higher-order moments of (4.3). This shows that the risk-sensitive performance index entails all the moments of (4.3). Note that when $\delta \to \infty$ ($\lambda \to 0$), the performance index (4.2) corresponds to the risk-neutral case considered in [71].

We define the following two sets of admissible controls:

$$\mathcal{U}_{1,i}^c = \{u_i : \; u_i(t) \in \sigma(x_i(s), s \le t, 1 \le i \le N), \tag{4.5}$$
$$\|x_i(T)\|^2 = o(T), \int_0^T \|x_i(t)\|^2 dt = O(T) \text{ a.s.}\}$$
$$\mathcal{U}_{1,i}^d = \{u_i : \; u_i(t) \in \sigma(x_i(s), s \le t), \tag{4.6}$$
$$\|x_i(T)\|^2 = o(T), \int_0^T \|x_i(t)\|^2 dt = O(T) \text{ a.s.}\},$$

where $\sigma(x_i(s), s \le t, 1 \le i \le N)$ is the σ-algebra generated by $x_i(s)$ for $s \le t$ and for all i, $1 \le i \le N$, and $\sigma(x_i(s), s \le t)$ is the σ-algebra generated by $x_i(s)$ for $s \le t$. It should be noted that the admissible control in (4.5) is centralized in terms of states of all agents, while the control in (4.6) is decentralized since it is associated with the local state information. Since $\mathcal{U}_{1,i}^d \subseteq \mathcal{U}_{1,i}^c$, we accordingly say that $\mathcal{U}_{1,i}^c$ is a set of admissible centralized controllers and $\mathcal{U}_{1,i}^d$ a set of admissible decentralized controllers for **P1**.

Now, under the admissible control sets defined above, **P1** is equivalent to characterizing an individual optimal control strategy that forms a Nash equilibrium (or ϵ-Nash equilibrium). The definition is given as follows.

Definition 4.1. *The set of controllers, $\{u_i^* \in \mathcal{U}_{1,i}^c,\ 1 \le i \le N\}$, constitutes an ϵ-Nash equilibrium with respect to the cost functions $\{J_{1,i}^N,\ 1 \le i \le N\}$, if there exists $\epsilon \ge 0$ such that for any i, $1 \le i \le N$,*

$$J_{1,i}^N(u_i^*, u_{-i}^*) \le \inf_{u_i \in \mathcal{U}_{1,i}^c} J_{1,i}^N(u_i, u_{-i}^*) + \epsilon.$$

If ϵ is zero, then $\{u_i^* \in \mathcal{U}_{1,i}^c,\ 1 \le i \le N\}$ is a Nash equilibrium with respect to $\{J_{1,i}^N,\ 1 \le i \le N\}$.

Note that in the definition, the Nash strategies, $\{u_i^* \in \mathcal{U}_{1,i}^c,\ 1 \le i \le N\}$, and the infimization on the RHS are not necessarily restricted to the decentralized set (4.6).

In Section 4.3, we characterize an ϵ-Nash equilibrium for **P1** where the corresponding strategy of each agent, say u_i^*, is decentralized, i.e. $u_i^* \in \mathcal{U}_{1,i}^d$ for all i. We also show that when the number of agents is large, $\{u_i^* \in \mathcal{U}_{1,i}^d,\ 1 \le i \le N\}$ is a Nash equilibrium.

4.2.2 LQ Robust Mean Field Games (P2)

We now consider the following SDE for agent i, $1 \le i \le N$:

$$dx_i = (A(\theta_i)x_i + B(\theta_i)u_i + D(\theta_i)v_i)dt + \sqrt{\mu}D(\theta_i)dW_i(t), \qquad (4.7)$$

where $x_i(0) = \bar{x}_i$ and $v_i \in \mathbb{R}^p$ is an auxiliary decision variable (*disturbance input*). The other parameters satisfy Assumption 4.1.

Similarly, we define the sets of admissible centralized and decentralized controllers for **P2** by

$$\begin{aligned}
\mathcal{U}_{2,i}^c = \{u_i :\ & u_i(t) \in \sigma(x_i(s), s \le t, 1 \le i \le N), \\
& \mathbb{E}\{\|x_i(T)\|^2\} = o(T), \mathbb{E}\{\int_0^T \|x_i(t)\|^2 dt\} = O(T)\ \} \\
\mathcal{U}_{2,i}^d = \{u_i :\ & u_i(t) \in \sigma(x_i(s), s \le t), \\
& \mathbb{E}\{\|x_i(T)\|^2\} = o(T), \mathbb{E}\{\int_0^T \|x_i(t)\|^2 dt\} = O(T)\ \}.
\end{aligned}$$

In this second problem, **P2**, each agent seeks a controller that minimizes the following worst-case performance index (defined for agent i):

$$\mathbf{P2}: \ J_{2,i}^N(u_i, u_{-i}) = \sup_{v_i \in \mathcal{V}_i} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\phi_i^2(x, f_N, u, v)\} \tag{4.8}$$

$$\phi_i^2(x, f_N, u, v) := \phi_i^1(x, f_N, u) - \int_0^T \gamma^2 \|v_i(t)\|^2 dt, \tag{4.9}$$

where $\mathcal{V}_i$ is the class of disturbances such that each $v_i(t)$ is adapted to $\sigma(x_i(s), s \le t, 1 \le i \le N)$, and $J_{2,i}^N(u_i, u_{-i}) < \infty$ for all $u_i \in \mathcal{U}_{2,i}^c$.

It should be noted that each agent is coupled via the mean field term (4.4). The disturbance v_i in (4.7) can be viewed as a fictitious player (or adversary) of agent i, which determines the worst-case risk-neutral performance index of agent i; therefore, the actual maximizing one in (4.7) for (4.8) will, in general, be different for different i. Notice also that this is not a $2N$-agent game, but is still an N-agent game with respect to (4.8). Now, viewed as an N-agent game, with the individual worst-case performance index (4.8), the two equilibrium concepts in Definition 4.1 are well-defined with respect to $\{J_{2,i}^N, \ 1 \le i \le N\}$ under $\mathcal{U}_{2,i}^c$.

In Section 4.5, we carry out the same analysis for **P2** as we do in Section 4.3 for **P1**. We then show in Section 4.6 that **P1** and **P2** are *partially equivalent* in the sense that the decentralized Nash strategies of the agents are similar, but are determined by different auxiliary systems that lead to the decentralized form.

4.3 LQ Risk-Sensitive Mean Field Games

In this section, we solve **P1** via risk-sensitive mean field control theory.

4.3.1 Risk-sensitive Optimal Control

This section solves a single LQ risk-sensitive optimal control problem.

We consider the following SDE:

$$dx = (Ax + Bu)dt + \sqrt{\mu} D dW(t), \ x(0) = \bar{x}, \tag{4.10}$$

and the risk-sensitive cost function

$$\bar{J}(u,g) = \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\bar{\phi}^1(x,g,u)}\} \tag{4.11}$$

$$\bar{\phi}^1(x,g,u) := \int_0^T \|x(t)-g(t)\|_Q^2 + \|u(t)\|_R^2 dt, \tag{4.12}$$

where $g \in \mathcal{C}_n^b$. The optimal control problem (4.11) can be seen as a robust tracking problem with respect to the given reference signal g [106]. We have the following result, where we suppress the subscript i and the parameter θ.

Proposition 4.1. *Consider the risk-sensitive control problem (4.11) with (4.10). Suppose that (A,B) is controllable and $(A,Q^{1/2})$ is observable. Suppose that for a fixed $\gamma := \sqrt{\delta/2\mu} > 0$, there is a matrix $P > 0$ that solves the following generalized algebraic Riccati equation (GARE):*

$$A^TP + PA + Q - P(BR^{-1}B^T - \frac{1}{\gamma^2}DD^T)P = 0. \tag{4.13}$$

Then,

(i) $H := A - BR^{-1}B^TP + \frac{1}{\gamma^2}DD^TP$ and $G := A - BR^{-1}B^TP$ are Hurwitz.

(ii) The optimal decentralized controller that minimizes (4.11) is given by

$$\bar{u}(t) = -R^{-1}B^TPx(t) - R^{-1}B^Ts(t), \tag{4.14}$$

where $s(t)$ satisfies the following differential equation:

$$\frac{ds(t)}{dt} = -H^Ts(t) + Qg(t), \tag{4.15}$$

with initial condition $s(0) = -\int_0^\infty e^{H^T\sigma}Qg(\sigma)d\sigma$.

(iii) The closed-loop system (4.10) with the optimal decentralized controller in (4.14) satisfies $\|x(T)\|^2 = o(T)$ and $\int_0^T \|x(t)\|^2 dt = O(T)$ almost surely.

(iv) The differential equation (4.15) has a unique solution in $\mathcal{C}_n^b$, which is given by

$$s(t) = -\int_t^\infty e^{-H^T(t-\sigma)}Qg(\sigma)d\sigma. \tag{4.16}$$

(v) *The minimum cost is*

$$\bar{J}(\bar{u}, g) = \limsup_{T\to\infty} \frac{1}{T}\int_0^T q(\tau)d\tau + \mu\,\mathrm{Tr}(PDD^T), \tag{4.17}$$

where $q(\tau) = g^T(\tau)Qg(\tau) - s^T(\tau)BR^{-1}B^Ts(\tau) + \gamma^{-2}s^T(\tau)DD^Ts(\tau)$.

Proof. See Appendix C.2. □

Remark 4.3. (i) *Define*

$$\gamma^* := \inf\{\gamma > 0 \ : \ P > 0,\ P \text{ solves (4.13)}\}. \tag{4.18}$$

Then by definition, for any finite $\gamma > \gamma^*$, *the GARE in (4.13) admits a solution* P. *In this case,* γ^* *is known as an optimum disturbance attenuation level, and any* $\gamma > \gamma^*$ *determines the level of the disturbance attenuation [106]. Moreover, given the system and the cost matrices,* γ^* *can be computed by transforming the GARE into the form of a linear matrix inequality (LMI) [115].*

(ii) *For any finite* $\gamma > \gamma^*$, *since* H *and* G *are Hurwitz, there exist positive constants* $\rho > 0$ *and* $\eta > 0$ *such that* $\|e^{Ht}\| \le \rho e^{-\eta t}$ *for all* $t \ge 0$. *The same holds for* G.

In the rest of the section, from Proposition 4.1, the triplet (A, B, D) is replaced by the ordered triplet $(A(\theta_i), B(\theta_i), D(\theta_i))$; accordingly the solution of the GARE in (4.13) is denoted by $P(\theta_i) > 0$. In that sense, $s_i(t)$ is the bias term of agent i with $(A(\theta_i), B(\theta_i), D(\theta_i), P(\theta_i))$. For notational convenience, we use the subscript i when $\theta = \theta_i$.

4.3.2 Mean Field Analysis

We now construct a mean field system for **P1** to characterize the best approximated mass behavior to the actual mass behavior (4.4) when N is large.

Let $\bar{x}_\theta(t) = \mathbb{E}\{x_\theta(t)\}$. By substituting the optimal decentralized controller (4.14) and the bias term (4.16) into (4.1), and taking expectation, $\bar{x}_\theta(t)$ can be written as

$$\bar{x}_\theta(t) = e^{G(\theta)t}x + \int_0^t e^{G(\theta)(t-\tau)}\bar{U}(\theta)\left(\int_\tau^\infty e^{-H(\theta)^T(\tau-\sigma)}Qg(\sigma)d\sigma\right)d\tau, \tag{4.19}$$

where $\bar{U}(\theta) = B(\theta)R^{-1}B^T(\theta)$. Then we can construct the following auxiliary system under Assumption 4.1:

$$\mathcal{T}(g)(t) := \int_{(\theta,x)\in\Theta\times X} \bar{x}_\theta(t)dF(\theta,x). \tag{4.20}$$

The operator (4.20) is a function of g, which captures the average behavior of all agents within the system parameter space Θ and the initial conditions when N is large; hence it must be consistent with the mass behavior (4.4) under (4.14). We call (4.20) the mean field system for **P1**.

Assumption 4.2. *(a) $(A(\theta), B(\theta))$ and $(A(\theta), Q^{1/2})$ are controllable and observable, respectively, for all $\theta \in \Theta$.*

(b) γ_θ^* is finite and $\gamma > \gamma_\theta^*$ for all $\theta \in \Theta$, where $\gamma_\theta^* := \inf\{\gamma > 0 : P(\theta) > 0,\ P(\theta)$ solves (4.13)$\}$.

(c) We have

$$\|R^{-1}\|\|Q\| \int_{\theta\in\Theta} \|B(\theta)\|^2 \Big(\int_0^\infty \|e^{G(\theta)\tau}\|d\tau\Big)\Big(\int_0^\infty \|e^{H(\theta)\tau}\|d\tau\Big)dF(\theta) < 1.$$

Remark 4.4. *For one-dimensional agent systems, by using the GARE, Assumption 4.2(c) can be simplified as follows*

$$\int_{\theta\in\Theta} \frac{QB^2(\theta)}{A(\theta)^2R + B(\theta)^2Q + \frac{1}{\gamma^2}A(\theta)D(\theta)^2P(\theta)R}dF(\theta) < 1.$$

As $\gamma \to \infty$, the above condition coincides with [71, Equation (18)].

We have the following result.

Theorem 4.1. *Suppose that Assumptions 4.1 and 4.2 hold. Then there is a unique $g^* \in \mathcal{C}_n^b$ such that $g^* = \mathcal{T}(g^*)$.*

Proof. From Lemma C.2(i) in Appendix C.1, the operator $\mathcal{T}(x) \in \mathcal{C}_n^b$ for any $x \in \mathcal{C}_n^b$, where $\mathcal{C}_n^b$ is a Banach space. Then for any $x, y \in \mathcal{C}_n^b$, we have

$$\begin{aligned}\|\mathcal{T}(x) - \mathcal{T}(y)\|_\infty \leq \|x-y\|_\infty \|R^{-1}\|\|Q\| \int_{\theta\in\Theta} \|B(\theta)\|^2 \\ \times \Big(\int_0^\infty \|e^{G(\theta)\tau}\|d\tau\Big)\Big(\int_0^\infty \|e^{H(\theta)\tau}\|d\tau\Big)dF(\theta).\end{aligned}$$

By Assumption 4.2 and the contraction mapping theorem [116], we have the desired result. $\square$

Remark 4.5. *Under Assumption 4.2(c), the infinite-dimensional fixed point problem in Theorem 4.1 can be solved via the Banach successive approximation method [116] or the policy iteration method [70].*

The optimal decentralized controller for agent i with g^* can be written as

$$u_i^*(t) = -R^{-1}B_i^T P_i x_i^*(t) - R^{-1}B_i^T s_i(t), \tag{4.21}$$

where $s_i(t)$ is determined by g^* in Theorem 4.1. Due to Proposition 4.1(iii), $u_i^* \in \mathcal{U}_{1,i}^d$. We denote the closed-loop system and the mass behavior under the optimal decentralized controller in (4.21) by $x_i^*(t)$ and f_N^*, respectively.

4.3.3 Closed-loop System Analysis

The following result shows the closed-loop system stability in the time-average sense.

Proposition 4.2. *Suppose that Assumptions 4.1 and 4.2 hold. Then,*

$$\sup_{N\geq 1} \max_{1\leq i\leq N} \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}[\int_0^T \|x_i^*(t)\|_Q^2 + \|u_i^*(t)\|_R^2 dt]}\} < \infty.$$

Proof. Note that

$$\begin{aligned}
\|x_i^*(t)\|_Q^2 \leq C_1\|e^{G_i t}x_i(0)\|_Q^2 + C_2\bigg\| \int_0^t e^{G_i(t-\tau)} B_i R^{-1} B_i^T s_i(\tau) d\tau \bigg\|_Q^2 \\
+ C_3 \bigg\| \sqrt{\mu} \int_0^t e^{G_i(t-\tau)} D_i dW_i(\tau) \bigg\|_Q^2 \\
\triangleq C_1 z_1(t) + C_2 z_2(t) + C_3 z_3(t),
\end{aligned}$$

and $\|u_i^*(t)\|_R^2 \leq C_4(\|x_i(t)\|^2 + \|s_i(t)\|^2)$; therefore, since $s_i \in \mathcal{C}_n^b$ for all i, we can choose a constant $C > 0$ independent of N such that

$$\|x_i^*(t)\|_Q^2 + \|u_i^*(t)\|_R^2 \leq C(z_1(t) + z_2(t) + z_3(t)).$$

Hence,

$$\limsup_{T\to\infty}\frac{\delta}{T}\log\mathbb{E}\{e^{\frac{1}{\delta}[\int_0^T \|x_i^*(t)\|_Q^2+\|u_i^*(t)\|_R^2 dt]}\}$$
$$\leq \limsup_{T\to\infty}\frac{\delta}{T}\log\mathbb{E}\{e^{\frac{1}{\delta}\int_0^T Cz_1(t)+Cz_2(t)+Cz_3(t)dt}\}$$
$$=\limsup_{T\to\infty}\Big[\frac{C}{T}\int_0^T z_1(t)+z_2(t)dt+\frac{\delta}{T}\log\mathbb{E}\{e^{\frac{C}{\delta}\int_0^T z_3(t)dt}\}\Big],$$

where the equality follows since z_1 and z_2 are deterministic processes.

In view of Lemma C.3 in Appendix C.1, $\limsup_{T\to\infty}\frac{C}{T}\int_0^T z_1(t)+z_2(t)dt<\infty$ that is independent of N due to the compactness of Θ. Let $\bar{z}_3(t)=\sqrt{\mu}\int_0^t e^{G_i(t-\tau)}D_i dW_i(\tau)$. Then by definition $z_3(t)=\|\bar{z}_3(t)\|_Q^2$, and $\bar{z}_3$ satisfies the following SDE:

$$d\bar{z}_3(t)=G_i\bar{z}_3(t)dt+\sqrt{\mu}D_i dW_i(t),\ \bar{z}_3(0)=0,$$

where G_i is Hurwitz. We introduce the auxiliary SDE

$$d\bar{z}_4(t)=G_i\bar{z}_4(t)dt+D_i w_i(t)dt+\sqrt{\mu}D_i dW_i(t),\ \bar{z}_4(0)=0.$$

Due to the connection between risk-sensitive and H^∞ control discussed in [83, 78, 79], we have

$$\limsup_{T\to\infty}\frac{\delta}{T}\log\mathbb{E}\{e^{\frac{C}{\delta}\int_0^T \|\bar{z}_3(t)\|_Q^2 dt}\}$$
$$=\sup_{w_i}\limsup_{T\to\infty}\frac{C}{T}\mathbb{E}\Big\{\int_0^T \|\bar{z}_4(t)\|_Q^2-\gamma^2\|w_i(t)\|^2 dt\Big\}.$$

Now, consider the following Riccati inequality

$$\begin{aligned}0&=A_i^T P_i+P_iA_i+Q-P_iB_iR^{-1}B_i^TP_i+\frac{1}{\gamma^2}P_iD_iD_i^TP_i\\&=G_i^TP_i+P_iG_i+Q+P_iB_iR^{-1}B_i^TP_i+\frac{1}{\gamma^2}P_iD_iD_i^TP_i\\&\geq G_i^TP_i+P_iG_i+Q+\frac{1}{\gamma^2}P_iD_iD_i^TP_i.\end{aligned}$$

Then in view of the above inequality, the KYP Lemma [117] and Example 2.2 in [78], and from the "completion of squares" method, the corresponding optimal solution is $w_i^*(t) = \frac{1}{\gamma^2} D^T P_i \bar{z}_4(t)$, which leads to

$$\sup_{w_i} \limsup_{T\to\infty} \frac{C}{T}\mathbb{E}\Big\{\int_0^T \|\bar{z}_4(t)\|_Q^2 - \gamma^2 \|w_i\|^2 dt\Big\}$$
$$\leq C\mu \operatorname{Tr}(P_i D_i D_i^T) \leq C\mu \sup_{\theta\in\Theta} \operatorname{Tr}(P(\theta)D(\theta)D^T(\theta)) < \infty,$$

where the last inequality follows from the compactness of Θ. This completes the proof. □

4.3.4 Consistency Analysis

We now show consistency between g^* and the actual mass behavior under (4.21) in the large population regime.

Theorem 4.2. *Suppose that Assumptions 4.1 and 4.2 hold. Then, the following hold:*

$$\text{(i)} \lim_{N\to\infty} \mathbb{E}\Big\{\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt\Big\} = 0,\ \forall T \geq 0$$
$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big\{\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt\Big\} = 0.$$

$$\text{(ii)} \lim_{N\to\infty} \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt = 0,\ \forall T \geq 0,\ a.s.$$
$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T} \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt = 0,\ a.s.$$

Proof. See Appendix C.2. □

Theorem 4.2 states that there is exact consistency between g^* and f_N^* in the mean-square sense as well as in the almost sure sense when the number of agents is arbitrarily large. We can also view this result as a law of large numbers with respect to N [71].

The following result establishes the concentration inequality of Theorem 4.2; the proof is given in Appendix C.2.

Corollary 4.1. *Let $X(N,T) := \int_0^T \|f_N^*(t) - g^*(t)\|^2 dt$. If Assumptions 4.1 and 4.2 hold, then* *(i) For each $c > 0$, there exists N' such that for all $N \geq N'$,*

$$\mathbb{P}(|X(N,T) - \mathbb{E}\{X(N,T)\}| \geq s) \leq 2e^{-2s^2/c^2},\ \forall T \geq 0.$$

(ii) For each $c(N) > 0$ with $N \geq 1$, there exists T' such that for all $T \geq T'$,

$$\mathbb{P}(|X(N,T) - \mathbb{E}\{X(N,T)\}| \geq sT) \leq 2e^{-2s^2/c(N)^2}.$$

In Corollary 4.1, with small c and $c(N)$, $X(N,T)$ is concentrated around its mean value with high probability. Note that in Corollary 4.1(ii), $c(N)$ can be chosen to be arbitrarily small when N is large due to Theorem 4.2(ii).

4.3.5 Asymptotic Equilibrium Analysis: The ϵ-Nash Equilibrium

This section characterizes an ϵ-Nash equilibrium for **P1**. We first introduce two cost functions that are related to admissible centralized and decentralized controllers:

$$J_{1,i}^N(u_i^*, u_{-i}^*) = \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\phi_i^1(x^*, f_N^*, u^*)}\}$$
$$J_{1,i}^N(u_i, u_{-i}^*) = \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\phi_i^1(x, f_N^{-i^*}, u)}\},$$

where $\phi_i^1(x^*, f_N^*, u^*)$ is (4.3) when all agents use the optimal decentralized controller (4.21) and $\phi_i^1(x, f_N^{-i^*}, u)$ is (4.3) when agent i is under the full state feedback controller $u_i \in \mathcal{U}_{1,i}^c$ (i.e., $x_i(t) := x_i(t)|_{u_i(t)}$), while other agents are still under the optimal decentralized controller (4.21).

We now state the main result, whose proof is given in Appendix C.2.

Theorem 4.3. *Suppose that Assumptions 4.1 and 4.2 hold. Then the set of the optimal decentralized controllers in (4.21), $\{u_i^* :\ 1 \leq i \leq N\}$, constitutes an ϵ-Nash equilibrium for **P1**. That is, for any i, $1 \leq i \leq N$, there exists $\epsilon_N \geq 0$ such that*

$$J_{1,i}^N(u_i^*, u_{-i}^*) \leq \inf_{u_i \in \mathcal{U}_{1,i}^c} J_{1,i}^N(u_i, u_{-i}^*) + \epsilon_N,$$

where $\epsilon_N \to 0$ as $N \to \infty$.

Remark 4.6. *Note that given g^*, **P1** is transformed into N independent risk-sensitive optimal control problems, and the resulting optimal controllers constitute an ϵ-Nash equilibrium.*

Remark 4.7. *As discussed in Remark 4.3(i), the individual Nash strategies have the disturbance attenuation properties with γ. This shows that the ϵ-Nash equilibrium features robustness, which will be demonstrated by simulations in Section 4.7.*

4.4 Limiting Behaviors of LQ Risk-Sensitive Mean Field Games

This section discusses limiting behaviors of **P1** with respect to two different design parameters. The first one is known as the large deviation limit (small noise limit) and the second one is the risk-neutral limit.

4.4.1 Small Noise Limit

We consider the small noise limit of **P1** in which the noise intensity parameter μ decreases to zero. If we take $\mu \to 0$, then the SDE for agent i now becomes the following ordinary differential equation:

$$\frac{dx_i(t)}{dt} = A(\theta_i)x + B(\theta_i)u_i.$$

Moreover, the risk-sensitive objective function in (4.2) heavily penalizes the large deviation of (4.3) when δ also decreases such that $\gamma = \sqrt{\delta/2\mu}$ is fixed and positive. In this case, the results in Proposition 4.1 and Theorems 4.1-4.3 are valid, since γ still remains as the same value. Specifically, under this limit, the set of N optimal controllers in (4.21) constitutes an ϵ-Nash equilibrium for **P1**.

4.4.2 Risk-neutral Limit

Now, consider the case when $\delta \to \infty$ for a fixed μ. As mentioned in Remark 4.2, under this limit, the risk-sensitive cost function (4.2) is equivalent to the risk-neutral case considered in [71]. This imposes a smaller weight on the large deviation of (4.3). We have the following result.

Proposition 4.3. *Consider* ***P1*** *with $\delta \to \infty$ for a fixed $\mu > 0$. If Assumptions 4.1 and 4.2 hold, then*

(i) The optimal decentralized controller for each agent can be written as

$$u_i^*(t) = -R^{-1}B_i^T Z_i x_i(t) - R^{-1}B_i^T r_i(t),$$

where $Z_i \geq 0$ is a solution of the following ARE:

$$A_i^T Z_i + Z_i A_i + Q - Z_i B_i R^{-1} B_i^T Z_i = 0,$$

and $r_i(t)$ is

$$r_i(t) = -\int_t^\infty e^{-F_i^T(t-\sigma)} Q g(\sigma) d\sigma,$$

where $F_i = A_i - B_i R^{-1} B_i^T Z_i$ is Hurwitz.

(ii) The set of N-decentralized controllers in (i) constitutes an ϵ-Nash equilibrium for ***P1****. Moreover, ϵ can be made arbitrarily small by picking N arbitrarily large.*

Proposition 4.3 implies that **P1** has the same limiting behavior of δ as the case of the LQ risk-sensitive control problem discussed in [75, 78, 106].

4.5 LQ Robust Mean Field Games

In this section, we solve **P2** formulated in Section 4.2.2 via worst-case mean field control theory.

4.5.1 The LQ Stochastic Zero-sum Differential Game

This section considers the LQ stochastic zero-sum differential game by replacing the mass behavior (4.4) with an arbitrary function $h \in \mathcal{C}_n^b$.

We consider the following SDE (from (4.7)):

$$dx = (Ax + Bu + Dv)dt + \sqrt{\mu} D dW(t),\ x(0) = \bar{x}. \tag{4.22}$$

The performance index is given by

$$\bar{J}(u, v, h) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\bar{\phi}^2(x, h, u, v)\} \tag{4.23}$$

$$\bar{\phi}^2(x, h, u, v) := \bar{\phi}^1(x, h, u) - \int_0^T \gamma^2 \|v(t)\|^2 dt. \tag{4.24}$$

We then have the following result:

Proposition 4.4. *Consider the LQ stochastic zero-sum differential game (4.22)-(4.23). Suppose that (A, B) is stabilizable and $(A, Q^{1/2})$ is detectable. Suppose that for a fixed $\gamma > 0$, there is a matrix $P \geq 0$ that solves the GARE (4.13). Then,*

(i) $H := A - BR^{-1}B^TP + \frac{1}{\gamma^2}DD^TP$ *and* $G := A - BR^{-1}B^TP$ *are Hurwitz.*

(ii) The optimal decentralized controller and the worst-case disturbance that constitute a saddle-point equilibrium[1] *for (4.23) are given by*

$$\bar{u}(t) = -R^{-1}B^TPx(t) - R^{-1}B^Ts(t) \tag{4.25}$$

$$\bar{v}(t) = \gamma^{-2}D^TPx(t) + \gamma^{-2}D^Ts(t), \tag{4.26}$$

where $s(t)$ satisfies the differential equation (4.15) with $g(t)$ replaced by $h(t)$.

(iii) The closed-loop system (4.22) with the optimal controller (4.25) and the worst-case disturbance (4.26) satisfies $\mathbb{E}\{\|x(T)\|^2\} = o(T)$ *and* $\mathbb{E}\{\int_0^T \|x(t)\|^2 dt\} = O(T)$.

(iv) $s(t)$ in (ii) has a unique solution in $\mathcal{C}_n^b$, which can be written as (4.16) with $g(t)$ replaced by $h(t)$.

(v) The saddle-point value is (4.17) with $g(t)$ replaced by $h(t)$.

Proof. See Appendix C.2. □

[1]See [61] or [106] for the definition of saddle-point equilibrium.

It should be mentioned that the optimal decentralized control laws of the individual agents for **P1** and **P2** are identical. This is to be expected, since the corresponding optimal control problems are equivalent in the sense that they share the same controller, and the minimum cost and the saddle-point value are identical [106, 78]. We also note that in Proposition 4.4, γ is a free variable.

Remark 4.8. *Just like the optimal control problem in Section 4.3.1, γ is the disturbance attenuation parameter that measures robustness of the closed-loop system [106]. Also, as $\gamma \to \infty$, (4.23) coincides with the LQ optimal control problem in [71].*

4.5.2 Mean Field Analysis

We now construct a mean field system for **P2**. Again, let $\bar{x}_\theta := \mathbb{E}\{x_\theta(t)\}$. By applying (4.25), (4.26), and (4.16) with $g(t)$ replaced by $h(t)$, $\bar{x}_\theta$ can be written as

$$\bar{x}_\theta(t) = e^{H(\theta)t}x + \int_0^t e^{H(\theta)(t-\tau)}U(\theta)\left(\int_\tau^\infty e^{-H^T(\theta)(\tau-\sigma)}Qh(\sigma)d\sigma\right), \quad (4.27)$$

where $U(\theta) = B(\theta)R^{-1}B^T(\theta) - \gamma^{-2}D(\theta)D^T(\theta)$. The mean field system for **P2** can be written as

$$\mathcal{L}(h)(t) := \int_{(\theta,x)\in\Theta\times X} \bar{x}_\theta(t)dF(\theta,x), \quad (4.28)$$

where $F(\theta,x)$ is the distribution function given in Assumption 4.1(e). Then, by following the argument as in Section 4.3.2, if the operator (4.28) is a contraction, then we have the fixed point. Notice that in general, its fixed point, say h^*, will not be identical to g^* in Theorem 4.1, provided that they exist, since h^* is also determined by the worst disturbance.

Assumption 4.3. *(a) $(A(\theta), B(\theta))$ and $(A(\theta), Q^{1/2})$ are stabilizable and detectable, respectively, for all $\theta \in \Theta$.*

(b) γ_θ^* is finite and $\gamma > \gamma_\theta^*$ for all $\theta \in \Theta$, where γ_θ^* is defined in Assumption 4.2.

(c) We have

$$\|Q\| \int_{\theta\in\Theta} \Big(\int_0^\infty \|e^{H(\theta)t}\|^2 dt\Big)^2 (\|B(\theta)\|^2\|R^{-1}\| + \gamma^{-2}\|D(\theta)\|^2) dF(\theta) < 1.$$

Remark 4.9. *As $\gamma \to \infty$, Assumption 4.3(c) becomes equivalent to that in [71].*

The following result is the counterpart of Theorem 4.1 in the present case.

Theorem 4.4. *Suppose that Assumptions 4.1 and 4.3 hold. Then there is a unique $h^* \in \mathcal{C}_n^b$ such that $h^* = \mathcal{L}(h^*)$.*

Now, the optimal decentralized controller for agent i is

$$u_i^*(t) = -R^{-1}B_i^T P_i x_i^*(t) - R^{-1}B_i^T s_i(t), \tag{4.29}$$

and the worst-case disturbance for agent i is

$$v_i^*(t) = \gamma^{-2}D_i^T P_i x_i^*(t) + \gamma^{-2}D_i^T s_i(t), \tag{4.30}$$

where $s_i(t)$ is now dependent on h^* in Theorem 4.4. Note that due to Proposition 4.4(iii), $u_i^* \in \mathcal{U}_{2,i}^d$ with v_i^*. By a possible abuse of notation, we denote the closed-loop system and the mass behavior under (4.29) and (4.30) by $x_i^*(t)$ and $f_N^*(t)$, respectively.

4.5.3 Closed-loop System Analysis

The next result shows the closed-loop system stability.

Proposition 4.5. *Suppose that Assumptions 4.1 and 4.3 hold. Then,*

$$\sup_{N\geq 1} \max_{1\leq i\leq N} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big\{\int_0^T \|x_i^*(t)\|^2 + \|u_i^*(t)\|^2 dt\Big\} < \infty.$$

Proof. The result can be shown in a similar way to that in Lemma C.3 in Appendix C.1. □

4.5.4 Consistency Analysis

The following result is the counterpart of Theorem 4.2 in this case.

Theorem 4.5. *Suppose that Assumptions 4.1 and 4.3 hold. Then, the following hold:*

$$\text{(i)} \lim_{N\to\infty} \mathbb{E}\Big\{\int_0^T \|f_N^*(t) - h^*(t)\|^2 dt\Big\} = 0, \ \forall T \geq 0$$
$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big\{\int_0^T \|f_N^*(t) - h^*(t)\|^2 dt\Big\} = 0.$$

$$\text{(ii)} \lim_{N\to\infty} \int_0^T \|f_N^*(t) - h^*(t)\|^2 dt = 0, \ \forall T \geq 0, \ \textit{a.s.}$$
$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T} \int_0^T \|f_N^*(t) - h^*(t)\|^2 dt = 0, \ \textit{a.s.}$$

The next result is the counterpart of Corollary 4.1.

Corollary 4.2. *Let* $Y(N,T) := \int_0^T \|f_N^*(t) - h^*(t)\|^2 dt$. *If Assumptions 4.1 and 4.3 hold, then*

(i) For each $c > 0$, *there exists* N' *such that for all* $N \geq N'$,

$$\mathbb{P}(|Y(N,T) - \mathbb{E}\{Y(N,T)\}| \geq s) \leq 2e^{-2s^2/c^2}, \ \forall T \geq 0.$$

(ii) For each $c(N) \geq 0$ *with* $N \geq 1$, *there exists* T' *such that for all* $T \geq T'$,

$$\mathbb{P}(|Y(N,T) - \mathbb{E}\{Y(N,T)\}| \geq sT) \leq 2e^{-2s^2/c(N)^2}.$$

4.5.5 Asymptotic Equilibrium Analysis: The ϵ-Nash Equilibrium

This section characterizes an ϵ-Nash equilibrium for **P2**. As in Section 4.3.5, we introduce two cost functions that are related to centralized and decentralized controllers:

$$J_{2,i}^N(u_i^*, u_{-i}^*) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\phi_i^2(x^*, f_N^*, u^*, v^*)\}$$
$$J_{2,i}^N(u_i, u_{-i}^*) = \sup_{v_i\in\mathcal{V}_i} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\phi_i^2(x, f_N^{-i^*}, u, v)\},$$

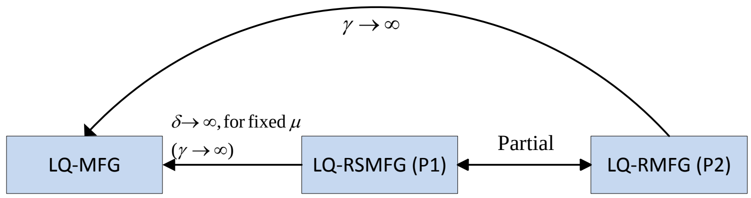


Figure 4.1: Relationship between **P1**, **P2**, and the risk-neutral LQ mean field game (LQ-MFG) with respect to the design parameters μ, δ, and γ.

where $\phi_i^2(x^*, f_N^*, u^*, v^*)$ is (4.9) when all the agents are under the optimal decentralized controller u_i^* in (4.29) and the worst-case disturbance v_i^* in (4.30), and $\phi_i^2(x, f_N^{-i^*}, u, v)$ is (4.9) when all the agents except i are under u_i^* and v_i^*, while agent i is under the centralized controller $u_i \in \mathcal{U}_{2,i}^c$ and $v_i \in \mathcal{V}_i$.

We now state the main result for **P2**.

Theorem 4.6. *Suppose that Assumptions 4.1 and 4.3 hold. Suppose that each agent adopts the corresponding worst-case disturbance in (4.30). Then, for any i, $1 \leq i \leq N$, there exists $\epsilon_N \geq 0$ such that*

$$J_{2,i}^N(u_i^*, u_{-i}^*) \leq \inf_{u_i \in \mathcal{U}_{2,i}^c} J_{2,i}^N(u_i, u_{-i}^*) + \epsilon_N,$$

where $\epsilon_N \to 0$ as $N \to \infty$.

Proof. See Appendix C.2. □

Remark 4.10. *Just like **P1**, given h^*, **P2** can be treated as N independent LQ stochastic zero-sum differential games, and the resulting N optimal controllers constitute an ϵ-Nash equilibrium.*

4.6 Discussion on Partial Equivalence and Limiting Behaviors of **P1** and **P2**

As we have seen in Sections 4.2-4.5, in **P1** and **P2**, the optimal controller for each agent that constitutes the ϵ-Nash equilibrium shares the same control law, but the bias term is determined by their respective mean field systems. This observation, and the limiting behaviors of **P1** discussed in Section 4.5 lead to the following conclusions.

There exists a *partial equivalence* between **P1** and **P2** in the sense that their respective approximated mass behaviors are different, although the individual optimal controllers that constitute ϵ-Nash equilibria share the same control laws. Specifically, the partial equivalence stems from the worst-case disturbance in **P2**. If $\gamma \to \infty$ (in **P1**, $\delta \to \infty$ for fixed μ), then the Nash equilibria of **P1** and **P2** become identical to that of the risk-neutral LQ mean field game (LQ-MFG) considered in [71], since the individual optimal controllers as well as the corresponding fixed points are identical to that in [71], as discussed in Proposition 4.3, and Remarks 4.8 and 4.9. Figure 4.1 summarizes this discussion.

It should be mentioned that a similar partial equivalence was discussed briefly in [73, Remark 6], and the result here sheds further light on that issue via the Nash certainty equivalence principle. In [73], it was noted that the control of fictitious player has to be included in the Fokker-Planck-Kolmogorov (FPK) equation of the robust mean field game that determines the density of the mass behavior, and thereby the mean field equilibrium solutions for the risk-sensitive and robust mean field games are not necessarily identical. The connection is that here the corresponding approximated mass behavior functions for **P1** and **P2**, g^* and h^*, resemble the FPK equation, and they are generally not identical due to the presence of the adversary in the latter.

4.7 Numerical Examples

In this section, two different numerical examples are considered. In each one, we compare **P1**, **P2**, and the risk-neutral case in [71] via their tracking performance and consistency between their respective approximated and actual mean field behaviors.

4.7.1 A System with Uniform Distribution

The first numerical example is the case when each agent's system parameter $A_i = \theta_i$ is an i.i.d. uniform random variable on the interval $[a, b]$ where $0 < a < b$. Also, $B = D = Q = R = 1$, $\mu = 2$, and the initial condition is 5. Note that under this setting, Assumption 4.1 holds where Assumption

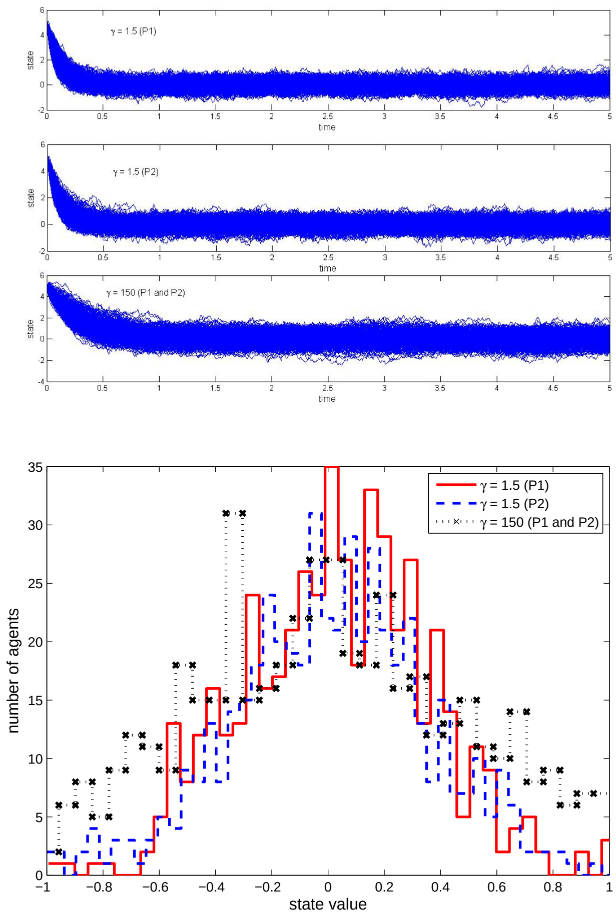


Figure 4.2: State trajectories of the scalar example when $N = 500$ (top) and the histogram (bottom) with $\gamma = 1.5$ and $\gamma = 150$.

4.1(e) holds due to Remark 4.1(ii). We obtain $P_i = \frac{\theta_i + \sqrt{\theta_i^2 + (1-\gamma^{-2})}}{(1-\gamma^{-2})} > 0$, which implies $\gamma_\theta^* = \gamma^* = 1$. Moreover, $H_i = -\sqrt{\theta_i^2 + (1-\gamma^{-2})} < 0$ and $G_i = \frac{-\gamma^{-2}\theta_i - \sqrt{\theta_i^2 + (1-\gamma^{-2})}}{(1-\gamma^{-2})} < 0$ for all i and any $\gamma > \gamma^* = 1$. Then the mean field systems for **P1** and **P2**, which are given in (4.20) and (4.28),

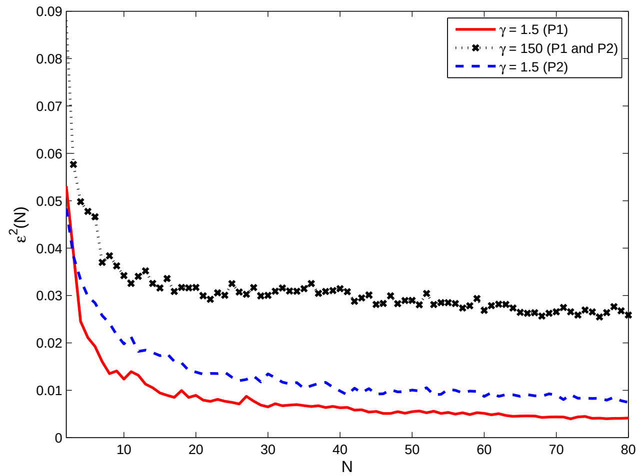


Figure 4.3: The approximation error, $\epsilon^2(N)$, with respect to N when $\gamma = 1.5$ and $\gamma = 150$.

respectively, can be written as

$$\mathcal{T}(g)(t) = \frac{1}{b-a}\int_{\theta\in[a,b]}\Big[5e^{G(\theta)t} + \int_0^t e^{G(\theta)(t-\tau)}\Big(\int_\tau^\infty e^{-H(\theta)(\tau-s)}g(s)ds\Big)d\tau\Big]d\theta \tag{4.31}$$

$$\mathcal{L}(h)(t) = \frac{1}{b-a}\int_{\theta\in[a,b]}\Big[5e^{H(\theta)t} + \int_0^t e^{H(\theta)(t-\tau)}(1-\gamma^{-2})\Big(\int_\tau^\infty e^{-H(\theta)(\tau-s)}h(s)ds\Big)d\tau\Big]d\theta. \tag{4.32}$$

Assume that $a = 2$, $b = 5$, and consider two cases; $\gamma = 1.5$ and $\gamma = 150$. We then need the contraction condition for (4.31) and (4.32) to characterize their fixed points. Note that since we are dealing with the heterogeneous agent case, the analytic method proposed in [70] is not applicable to obtain the fixed points.

It can be verified that Assumptions 4.2(c) and 4.3(c) hold; hence we have contraction. Then we use the Banach successive approximation method [116], by which the fixed point of (4.31) and (4.32) when $\gamma = 1.5$ are $g^*(t) = 5.086e^{-8.49t}$ and $h^*(t) = 5.1e^{-3.37t}$, respectively, where the exact values are obtained by the curve fitting algorithm. Note that due to the partial equivalence, $g^* \not\equiv h^*$, and $g^*, h^* \in \mathcal{C}_1^b$. The same method is used to obtain the fixed points of (4.31) and (4.32) when $\gamma = 150$.

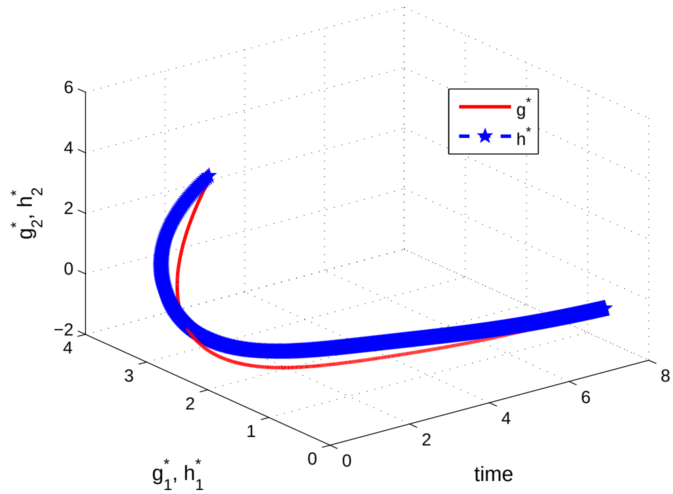


Figure 4.4: The approximated mass behavior, $g^* = (g_1^*, g_2^*)^T$ and $h^* = (h_1^*, h_2^*)^T$, of **P1** and **P2**, respectively, for the second order damping system when $\gamma = 3$.

Figure 4.2(top) shows the state trajectories of 500 agents for **P1** and **P2**. Each trajectory and therefore the equilibrium features robustness when $\gamma = 1.5$, which is a consequence of its disturbance attenuation property with γ as discussed earlier in the chapter. This is well illustrated in the state histogram in Fig. 4.2 (bottom). Each agent is more concentrated when $\gamma = 1.5$. Note that when $\gamma = 150$, the trajectories of **P1** and **P2** show the same behaviors, which is a consequence of their limiting behavior with respect to γ. The curve of the approximation performance, $\epsilon^2(N)$, with respect to N is shown in Fig. 4.3, where $\epsilon^2(N) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt\}$ (or h^*). Each plot shows convergence approximately at the rate of $O(1/N)$, and also shows that the curve with a smaller γ leads to a better approximation performance due to the disturbance attenuation property of **P1** and **P2**.

4.7.2 A Second-order Damping System

The next numerical example considers the transfer function of the following second-order damping system:

$$H_i(s) = \frac{1}{s^2 + 2s + \beta_i}, \tag{4.33}$$

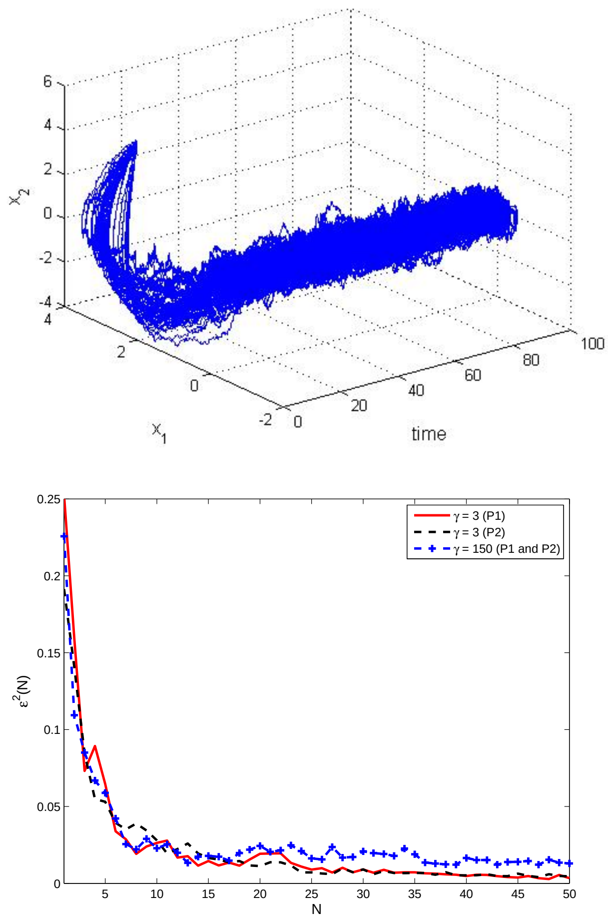


Figure 4.5: State trajectories of the second order damping system when $N = 50$ (top), and $\epsilon(N)$ with respect to N (bottom). Note that the equilibrium features robustness with $\gamma = 3$.

whose state space representation is

$$A_i = \begin{pmatrix} 0 & 1 \\ -\beta_i & -2 \end{pmatrix}, \ B_i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We assume $D = (0,1)^T$, $Q = I$, $R = 5$, $\mu = 0.5$, and the initial condition is taken to be $(2,5)^T$. Also, β_i is an i.i.d. uniform random variable on $[-4,4]$. Note that the pole of (4.33) is located at a radius of $\beta_i^{1/2}$ and at an angle of $\sin^{-1}(1/\beta_i^{1/2})$; hence, depending on realizations of β_i, the system can be stable or unstable.

By using the numerical integration method, it can be checked that the mean field systems for **P1** and **P2** are contraction when $\gamma = 3$. Figure

4.4 shows the 2-dimensional approximated mass behaviors, g^* and h^*, for the second-order system, which are obtained by the Banach successive approximation method. Figure 4.5 shows the trajectory of $N = 50$ and the approximation error, $\epsilon^2(N)$, with respect to N. Notice that better approximation performance is achieved when $\gamma = 3$ due to the inherent robustness property of **P1** and **P2**.

4.8 Conclusions

In this chapter, two classes of LQ mean field games have been considered; the linear-quadratic risk-sensitive mean field game (**P1**) and the linear-quadratic robust mean field game (**P2**). We have obtained ϵ-Nash equilibria for both **P1** and **P2**, where the individual Nash strategies are decentralized as a function of an agent's own state and the best approximated mass behavior function, where the latter can be determined off-line. Specifically, under the individual optimal decentralized controllers, the mass behavior collapses into the best approximated function that is obtained by fixed point analysis of the corresponding mean field system. We have shown that ϵ-Nash equilibria for **P1** and **P2** are partially equivalent in the sense that the Nash strategies share the same control laws, but are also characterized by different approximated mass behavior functions. This partial equivalence stems from the difference between their respective mean field systems due to the presence of the adversary in **P2**.

We have also studied the limiting behaviors of the ϵ-Nash equilibria in the large population regime as well as in the limit of the design parameters. In particular, we have shown that when the population size grows to infinity, ϵ-Nash equilibria become (exact) Nash equilibria. Moreover, when the disturbance attenuation parameter goes to infinity, ϵ-Nash equilibria of **P1** and **P2** become identical to that of the risk-neutral LQ mean field game considered in [71]. Note that such a relationship can be viewed as the mean field counterpart of what is observed in the one-agent case discussed in earlier literature, see [75, 106, 78, 79, 77]. Finally, we have shown that the ϵ-Nash equilibria for **P1** and **P2** feature robustness due to the risk-sensitivity of the former and the worst-case characteristic of the latter. This robustness property has been demonstrated by two different numerical examples.

CHAPTER 5

MEAN FIELD STACKELBERG DIFFERENTIAL GAMES

5.1 Introduction

In this chapter, we consider mean field Stackelberg differential games when there is one leader and a large number, say N, of followers. The leader *globally* dominates over the followers for the entire duration in the sense that before the start of the game he chooses and then announces his strategy to the N followers who play a Nash game. The N number of Nash followers choose their optimal strategies noncooperatively and simultaneously based on the leader's observed strategy. Note that the class of Stackelberg games with one leader and N number of Nash followers was studied (without the mean field framework) in [89, 118].

The information structures for both the leader and the followers are *adapted open-loop*,[1] that is, information on each agent's initial condition and filtration is available to each agent (the leader, of course, knows everything that the followers know). Moreover, in this setting, the followers are coupled with each other through the mean field term included in each follower's cost function, and are strongly influenced by the leader's strategy included in each agent's cost function and dynamics. From the leader's perspective, he is coupled with the N followers through the mean field term included in his cost function. We also consider the heterogeneous case of the followers with K distinct models, that is, follower i belongs to a finite model set $\mathcal{K} = \{1, 2, ..., K\}$.

Since there is a large number of followers, complexity issues arise from the mean field coupling term and heterogeneity of Nash followers. In addition, solving the leader's optimal control problem becomes complicated, since it depends on a large number of Nash followers. Hence, computing an exact

[1] The notion of (adapted) open-loop information structure for (stochastic) Stackelberg games can be found in [61, 93].

Stackelberg-Nash solution is cumbersome to say the least. To circumvent this difficulty, our approach in this chapter is to apply the stochastic mean field approximation to characterize the best estimate of the actual mean field behavior.

We first consider the mean field Nash game for the N followers given an arbitrary strategy of the leader. We solve a local optimal control problem of the followers with leader's control taken as an exogenous stochastic process. We characterize the best estimate of the actual mean field behavior that is dependent on the leader's arbitrary strategy. Note that the local optimal controller for the followers is decentralized, as it is a function of local information and the approximated mean field process (and also the leader's arbitrary strategy). We show that for each fixed strategy of the leader, the followers' optimal decentralized strategies lead to an ϵ-Nash equilibrium, where ϵ converges to zero as $N \to \infty$.

We then consider the leader's problem. The leader's local optimal control problem includes additional constraints induced by the mean field process determined by Nash followers, which is thus still hard to solve, but is much more tractable than the leader's original optimal control problem, since in the latter, the number of additional constraints depends on N. We obtain the leader's decentralized optimal controller as a function of his information and the mean field process. Unlike the follower's problem, the leader's problem may not have a unique optimal solution due to coupled forward-backward stochastic differential equations (FBSDEs) and the corresponding *nonsymmetric* Riccati differential equation (RDE). We identify a linear matrix inequality (LMI) condition under which solutions of the RDE as well as the FBSDEs exist. We finally show that the optimal decentralized controllers for the leader and the followers constitute an (ϵ_1, ϵ_2)-Stackelberg equilibrium for the original game, where ϵ_1 and ϵ_2 both converge to zero as $N \to \infty$. This implies that for a large number of followers, the impact of the followers on the leader collapses to the approximated mean field process, which reduces complexity of the original Stackelberg game.

Organization

The chapter is organized as follows. In Section 5.2, we formulate the problem, and discuss the difficulty of obtaining the solution to the original problem when the number of followers is large. Section 5.3 solves the follower's problem given an arbitrary strategy of the leader, and characterizes the mean field process. The leader's problem is discussed in Section 5.4, where we obtain an approximated Stackelberg equilibrium. Numerical examples are presented in Section 5.5. We end this chapter with the concluding remarks of Section 5.6.

5.2 LQ Mean Field Stackelberg Games

In this section, we formulate the problem of linear-quadratic (LQ) mean field Stackelberg games, and discuss the difficulty in obtaining the solution with a finite (but large) number of followers.

5.2.1 Problem Formulation

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, t \geq 0, \mathbb{P})$ be a natural complete filtered probability space augmented by all the $\mathbb{P}$ null sets in $\mathcal{F}$. We denote the set of all vector-valued (or matrix-valued) $\mathcal{F}_t$ adapted processes satisfying $\mathbb{E}\int_0^T \|\cdot\|^2 dt < \infty$ by $L^2_{\mathcal{F}}(0,T;\cdot)$.

We have one leader, $\mathcal{P}_0$, and N followers, $\{\mathcal{P}_i,\ 1 \leq i \leq N\}$. The leader and the followers have their own cost functions $\{J_i^N,\ 0 \leq i \leq N\}$ which each wants to minimize.

We consider the controlled stochastic differential equation (SDE) for the leader, $\mathcal{P}_0$,

$$dx_0(t) = [A_0 x_0(t) + B_0 u_0(t)]dt + D_0 dW_0(t), \tag{5.1}$$

where $x_0 \in \mathbb{R}^n$ is the state that captures the behavior of $\mathcal{P}_0$, $u_0 \in \mathbb{R}^p$ is the control of $\mathcal{P}_0$, and $\{W_0(t), t \geq 0\}$ is a q-dimensional Brownian motion. The SDE for the follower $\mathcal{P}_i$, $1 \leq i \leq N$, is given by

$$dx_i(t) = [A(\theta_i)x_i(t) + Bu_i(t) + Fu_0(t)]dt + DdW_i(t), \tag{5.2}$$

where $x_i \in \mathbb{R}^n$ is the state of follower $\mathcal{P}_i$, $u_i \in \mathbb{R}^p$ is the control of $\mathcal{P}_i$, and $\{W_i(t), t \geq 0\}$ is a q-dimensional Brownian motion. In (5.1) and (5.2), A_0, B_0, D_0, $A_i := A(\theta_i)$, B, D, and F are time-invariant matrices with appropriate dimensions. Let $\mathcal{F}_t$ be the σ-algebra generated by $\{x_i(0), 0 \leq i \leq N\}$ and $\{W_i(\tau), \tau \leq t, 0 \leq i \leq N\}$, i.e., $\mathcal{F}_t = \sigma(x_i(0), W_i(\tau), \tau \leq t, 0 \leq i \leq N)$. Let $\mathcal{F}_t^i = \sigma(x_i(0), W_i(\tau), \tau \leq t)$, for i, $0 \leq i \leq N$. We say that $\mathcal{F}_t^i$ is local information of agent i, $0 \leq i \leq N$, whereas $\mathcal{F}_t$ is the global (or centralized) information.

In (5.2), $\{\theta_i, 1 \leq i \leq N\}$ models the heterogeneity of the followers, which can be viewed as a sequence of dynamic parameters. For simplicity, we consider the case that only the system matrices of the followers, $A_i = A(\theta_i)$, $1 \leq i \leq N$, are different, and the analysis developed in this chapter can easily be extended to the case when other parameters are also dependent on θ_i. We assume that $\mathcal{P}_i$, $1 \leq i \leq N$, is of K distinct models. Specifically, let $\mathcal{K} = \{1, 2, ..., K\}$ and $\mathcal{N}_k = \{i : \theta_i = k, 1 \leq i \leq N\}$ for $k \in \mathcal{K}$. Then $N_k := |\mathcal{N}_k|$ where $|\mathcal{N}_k|$ is the cardinality of $\mathcal{N}_k$ capturing K distinct types of followers with $\{\theta_i, 1 \leq i \leq N\}$. In our heterogeneous model, we have $A_i = A_j$ if $i, j \in \mathcal{N}_k$, in which case we denote $A_k := A_i = A_j$. Since $N = \sum_{k=1}^{K} N_k$, the vector $\pi^N = (\pi_1^N, ..., \pi_K^N)$, where $\pi_k^N = N_k/N$, $k \in \mathcal{K}$, becomes the probability distribution on $\{\theta_i, 1 \leq i \leq N\}$.

We introduce the following assumption:

Assumption 5.1. *(a)* $\{x_i(0), 0 \leq i \leq N\}$ *are independent of each other, and are measurable on* $\mathcal{F}_0$. *Also,* $\mathbb{E}[x_0(0)] = \bar{x}_0$, $\mathbb{E}[x_i(0)] = \bar{x}$, *and* $\sup_{i \geq 0} \mathbb{E}[\|x_i(0)\|^2] \leq c < \infty$.

(b) $\{W_i(t), 0 \leq i \leq N\}$ *are independent of each other, which are also independent of* $\{x_i(0), 0 \leq i \leq N\}$.

(c) There exists a probability vector $\pi = (\pi_1, ..., \pi_K)$ *such that* $\lim_{N \to \infty} \pi^N = \pi$, *where* $0 < \pi_k \leq 1$ *for all* $k \in \mathcal{K}$ *with* $\sum_{k=1}^{K} \pi_k = 1$.

It should be mentioned that Assumption 5.1(c) implies that $N_k > 0$ for each $k \in \mathcal{K}$, and $N_k \to \infty$ when $N \to \infty$, that is, any type of the followers is not diminished when N goes to infinity. One example of Assumption 5.1(c) is $\mathcal{K} = \{1, 2\}$ with $N_1 = N_2 = N/2$, in which case $\pi = (0.5, 0.5)$. Finally, due to Assumptions 5.1(a) and 5.1(b), we have $\mathcal{F}_t = \vee_{i=0}^{N} \mathcal{F}_t^i$.

The performance index for $\mathcal{P}_0$ to be minimized is given by

$$J_0^N(u_0, u^N) = \mathbb{E} \int_0^T \Big[\|x_0(t) - H_0 x^N(t)\|_{Q_0}^2 + \|u_0(t)\|_{R_0}^2 \Big] dt, \tag{5.3}$$

where $Q_0 \geq 0$ and $R_0 > 0$. In (5.3), $x^N(t) = \frac{1}{N}\sum_{i=1}^N x_i(t)$ denotes the *mean field* term that captures the mass behavior of the followers. Note that in view of (5.3), the leader and the followers are coupled through the mean field term. The performance index for $\mathcal{P}_i$, $1 \leq i \leq N$, is given by

$$J_i^N(u_i, u_{-i}, u_0) = \mathbb{E} \int_0^T \Big[\|x_i(t) - H x^N(t)\|_Q^2 + \|u_i(t)\|_R^2 + 2u_i^T(t) L u_0(t) \Big] dt, \tag{5.4}$$

where $Q \geq 0$ and $R > 0$. Note that the followers are (weakly) coupled with each other through the mean field term x^N, but are strongly coupled with the leader's control u_0 included in (5.2) and (5.4).

The classes of admissible controls for $\mathcal{P}_0$ and $\mathcal{P}_i$ are defined as follows. Let $z \in L_{\mathcal{F}}^2(0, T; \mathbb{R}^n)$. We define $\mathcal{U}_0$ to be the set of $u_0 \in L_{\mathcal{F}}^2(0, T; \mathbb{R}^p)$ such that u_0 being a function of $y = (x_0, ..., x_N, z)$ is Lipschitz continuous with respect to y. Similarly, we define $\mathcal{U}_i(u_0)$ to be the set of $u_i \in L_{\mathcal{F}}^2(0, T; \mathbb{R}^p)$ such that given $u_0 \in \mathcal{U}_0$, u_i being a function of $y' = (x_1, ..., x_N, u_0, z)$ is Lipschitz continuous with respect to y' for all i. Note that under these two definitions, for any $u_0 \in \mathcal{U}_0$ and $u_i \in \mathcal{U}_i(u_0)$, there exists a unique (strong) solution to SDEs in (5.1) and (5.2) in $L_{\mathcal{F}}^2(0, T; \mathbb{R}^n)$ [119]. It should also be noted that in the definitions of $\mathcal{U}_0$ and $\mathcal{U}_i(u_0)$, the corresponding controllers are not restricted to be functions of the local information.

We next discuss the hierarchy of the (adapted) open-loop Stackelberg game under consideration. The leader, $\mathcal{P}_0$, holds a dominating position in the sense that he first decides on and announces his strategy u_0, and enforces on $\mathcal{P}_i$, $1 \leq i \leq N$; in this process, the leader takes into account the rational reactions of the followers. The N followers then respond by playing a Nash game under the leader's strategy. Note that in this framework, each player (leader and followers) knows his own system and cost parameters, and $\mathcal{P}_0$ also knows the cost and system parameters of the followers. Moreover, the followers know the leader's optimal strategy for the entire time-horizon ($[0, T]$), that is, the leader announces his strategy in advance, and ahead of the followers at the start of the game.

In this hierarchical setting, the main objective of this chapter is to obtain a Stackelberg equilibrium[2] for the corresponding game. However, since we are dealing with a large number of followers, two different issues of complexity arise. One is with regard to the Nash game of the followers, since it is hard (maybe impossible) to obtain the Nash equilibrium with an arbitrary strategy of the leader when N is large. Another one is with regard to the stochastic optimal control problem faced by the leader, as the number of additional constraints induced by the Nash followers grows with N. This point will be discussed in more detail in Section 5.2.2.

Now, to address these two issues of complexity, instead of seeking a Stackelberg solution, we seek an approximated Stackelberg equilibrium by characterizing the best estimate of the actual mean field behavior x^N when N is arbitrarily large, in view of which 1) the Nash game of the followers admits an ϵ-Nash equilibrium for any arbitrary open-loop strategy of the leader, and 2) the leader is able to characterize a unique optimal solution that leads to an approximated Stackelberg strategy. We will see that due to the mean field approximation, the number of additional constraints for the leader's optimization problem induced by the Nash followers will be independent of N, making it easier to solve than the leader's original optimization problem.

5.2.2 Discussion on the Direct Approach

We discuss the direct approach to the Stackelberg game formulated in Section 5.2.1. We do not provide an explicit expression of the Stackelberg equilibrium, but (informally) argue why the standard direct approach discussed in [61, 93] for games with a small number of players may not be applicable to solve the mean field Stackelberg game formulated in Section 5.2.1.

We first need to solve the Nash game of the followers for an arbitrary open-loop strategy of the leader. Suppose that each follower has access to the global state information. From [119], the Hamiltonian for follower i given

[2] See [61] for the definition of a Stackelberg equilibrium.

$u_0 \in \mathcal{U}_0$ can be written as (t argument is suppressed)

$$\begin{aligned}\mathcal{H}_i(x_i, u_i, s_i^N, g_i) = &-\Big[\|x_i - Hx^N\|_Q^2 + \|u_i\|_R^2 + 2u_i^T L u_0\Big] \\ &+ s_i^{N,T}(A_i x_i + Bu_i + Fu_0) + \mathrm{Tr}(g_i^T D).\end{aligned}$$

Then by the stochastic maximum principle (see [119]), the optimal solution[3] for follower i is the one that maximizes the above Hamiltonian:

$$u_i'(t) = R^{-1}B^T s_i^N(t) - R^{-1}Lu_0(t), \tag{5.5}$$

where the corresponding forward-backward SDE (FBSDE) is given by

$$\begin{cases} dx_i'(t) = \Big[A_i x_i'(t) + BR^{-1}B^T s_i^N(t) - (BR^{-1}L - F)u_0(t)\Big]dt + DdW_i(t) \\ ds_i^N(t) = \Big[-A_i^T s_i^N(t) + Qx_i'(t) - QHx^N(t) - (1/N)H^TQx_i'(t) \\ \qquad + (1/N)H^TQHx^N(t)\Big]dt + g_i(t)dW_i(t) \\ x_i'(0) = x_i(0),\ s_i^N(T) = 0. \end{cases} \tag{5.6}$$

As can be seen from (5.6), the dynamics for x_i' depend on the leader's open-loop strategy, and the adjoint process s_i^N is a function of the mean field term. Note that this optimality condition holds for any follower. Therefore, since each follower is coupled through the cost functionals only, if there is an $\mathcal{F}_t$-adapted solution of the FBSDE in (5.6) for all i, $1 \leq i \leq N$, then the set of u_i' in (5.5) leads to a Nash equilibrium of the followers given any arbitrary open-loop strategy of the leader [61]. In general, however, neither the existence of the solution of (5.6) nor its uniqueness is guaranteed, and even if a solution exists, it is hard to obtain, since (5.6) is highly coupled across followers. Moreover, it may not be realistic in many cases that each follower has access to the state information of the other followers. This corresponds to the complexity of the Nash game when the number of followers is large as mentioned in Section 5.2.1.

Now, to obtain a Stackelberg strategy for the leader, the leader is faced

[3]Note that in our case, the maximum principle is also sufficient, since $R > 0$ and $\mathcal{H}_i$ is concave in x_i for all i, see [119, Theorem 5.2].

with the following optimization problem:

$$\min_{u_0} J_0^N(u_0, u^N), \text{ s.t. } x_0,\ x_i',\ s_i^N,\ i = 1, 2, ..., N, \tag{5.7}$$

where x_0 is the state equation for the leader, and x_i' and s_i^N are from (5.6), which are additional constraints induced by Nash followers. Note that in the above optimization problem, the cost function $J_0^N(u_0, u^N)$ includes the mean field term $\frac{1}{N}\sum_{i=1}^N x_i'(t)$ which is dependent on u_0 by virtue of (5.6). Moreover, the number of additional constraints increases with N, which demonstrates the second complexity issue brought up in Section 5.2.1. Hence, the above heuristic argument informally shows that the standard approach may not be applicable to obtain a Nash-Stackelberg equilibrium, and strengthens our motivation for studying the problem within the mean field theory framework.

5.3 Mean Field Nash Games for the N Followers

This section considers the mean field Nash game for the N followers under an arbitrary strategy of the leader, $u_0 \in \mathcal{U}_0$. Since the followers are coupled through only the mean field term, under the mean field approximation, each follower actually faces a separate stochastic optimal control problem, which we discuss below.

5.3.1 Local Optimal Control Problem for $\mathcal{P}_i$

Consider the SDE for $\mathcal{P}_i$, $1 \le i \le N$:

$$dx_i(t) = [A(\theta_i)x_i(t) + Bu_i(t) + Fu_0(t)]dt + DdW_i(t), \tag{5.8}$$

where $\mathbb{E}[x_i(0)] = \bar{x}$ with $\mathbb{E}[\|x_i(0)\|^2] < \infty$. The cost function is given by

$$\bar{J}_i(u_i) = \mathbb{E}\int_0^T \Big[\|x_i(t) - Hz(t)\|_Q^2 + \|u_i(t)\|_R^2 + 2u_i^T(t)Lu_0(t)\Big]dt, \tag{5.9}$$

where $z(t) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)$ replaces $x^N(t)$ in (5.4), which can be viewed as the mass behavior when $N \to \infty$, whose explicit expression will be derived later in this section. Note that due to the hierarchical structure of the game under consideration, z will be determined by an arbitrary strategy of $\mathcal{P}_0$. Moreover, the individual impact of each follower on z may be negligible if z captures the mass behavior.

The optimal control problem (5.9) can be viewed as a standard stochastic LQ problem by treating z and u_0 as exogenous signals. We have the following result, whose proof follows from the standard stochastic optimal control problem, see [120, 121, 119].

Lemma 5.1. *Given $z \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^n)$ and $u_0 \in \mathcal{U}_0$, consider the local optimal control problem for $\mathcal{P}_i$, $1 \leq i \leq N$, in (5.9). There exists a unique optimal controller $u_i^* \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^q)$. Moreover, $(x_i^*, u_i^*) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n+p})$ is the corresponding optimal solution if and only if*

$$u_i^*(t) = R^{-1}B^T p_i(t) - R^{-1}Lu_0(t), \tag{5.10}$$

where

$$\begin{cases} dx_i^*(t) = \Big[A_i x_i^*(t) + BR^{-1}B^T p_i(t) - (BR^{-1}L - F)u_0(t)\Big]dt + DdW_i(t) \\ dp_i(t) = \Big[-A_i^T p_i(t) + Q(x_i^*(t) - Hz(t))\Big]dt + r_i(t)dW_i(t) \\ x_i^*(0) = x_i(0), \; p_i(T) = 0, \end{cases} \tag{5.11}$$

where $(x_i^*, p_i, r_i) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{2n}, \mathbb{R}^{n\times q})$ is the solution to the forward-backward SDE (FBSDE) in (5.11). Finally, given z and u_0, (x_i^*, p_i, r_i) has a unique solution in $L^2_{\mathcal{F}}(0, T; \mathbb{R}^{2n}, \mathbb{R}^{n\times q})$.

We now obtain an equivalent state-feedback representation of the optimal controller and the corresponding optimal trajectory given in (5.10) and (5.11), respectively. Let $\mathbb{G}_i(t) := A_i - BR^{-1}B^T Z_i(t)$ and $p_i(t) = -Z_i(t)x_i^*(t) + \phi_i(t)$, where $\phi_i(T) = 0$ and

$$-\frac{dZ_i(t)}{dt} = A_i^T Z_i(t) + Z_i(t)A_i + Q - Z_i(t)BR^{-1}B^T Z_i(t), \; Z_i(T) = 0. \tag{5.12}$$

Notice that (5.12) is a symmetric Riccati differential equation (RDE) that arises in the standard LQ optimal control problem. We can show that

$$\begin{cases} dx_i^*(t) = \Big[\mathbb{G}_i(t)x_i^*(t) + BR^{-1}B^T\phi_i(t) - (BR^{-1}L - F)u_0(t)\Big]dt + DdW_i(t) \\ d\phi_i(t) = \Big[-\mathbb{G}_i^T(t)\phi_i(t) - QHz(t) - Z_i(t)(BR^{-1}L - F)u_0(t)\Big]dt \\ \qquad\qquad + (Z_i(t)D + r_i(t))dW_i(t) \\ x_i^*(0) = x_i(0),\ \phi_i(T) = 0, \end{cases} \tag{5.13}$$

where the corresponding optimal decentralized state feedback controller is given by

$$u_i^*(t) = -R^{-1}B^TZ(t)x_i^*(t) + R^{-1}B^T\phi_i(t) - R^{-1}Lu_0(t). \tag{5.14}$$

Note that ϕ_i is now decoupled from x_i^*. Since $R > 0$ and $Q \geq 0$, from the standard LQ optimal control theory (see [119]), the RDE in (5.12) admits a unique solution $Z_i(t) \geq 0$ for all $t \in [0, T]$ with $Z_i(T) = 0$. Moreover, we have $Z_k := Z_i \equiv Z_j$ if $i, j \in \mathcal{N}_k$ for any $k \in \mathcal{K}$. Therefore, since $\mathcal{K}$ is finite, $\sup_{k\in\mathcal{K},0\leq t\leq T} \|Z_k(t)\| \leq C < \infty$. This also implies that given z and u_0, there exists a unique $\mathcal{F}_t$ adapted solution to the FBSDE in Lemma 5.1 due to the affine structure of the transformation [121, Theorem 4.1]. The explicit expression of the solution of (5.13) (and therefore (5.11)) in integral form can be obtained by using [85, Lemma A.1].

Remark 5.1. *It should be noted that the local optimal controller for $\mathcal{P}_i$ is a function of its own information (including z) and an arbitrary open-loop strategy of the leader. In view of this, we call (5.10) or (5.14) the optimal decentralized controller for the follower.*

5.3.2 Stochastic Mean Field Approximation

This subsection considers the mean field approximation. We first apply the optimal decentralized controller (5.10) (or (5.14)) to N followers, and denote the corresponding state by x_i^*. From our K distinct heterogeneity model, let

$z_k^N(t) = \frac{1}{N_k}\sum_{i\in\mathcal{N}_k} x_i^*(t)$. Then

$$x^N(t) = \frac{1}{N}\sum_{i=1}^{N} x_i^*(t) = \frac{1}{N}\sum_{k=1}^{K}\sum_{i\in\mathcal{N}_k} x_i^*(t) = \frac{1}{N}\sum_{k=1}^{K} N_k z_k^N(t) = \sum_{k=1}^{K}\pi_k^N z_k^N(t),$$

where $\pi_k^N = N_k/N$ is the probability of the k-th model in $\mathcal{K}$ defined in Section 5.2.1.

Now, z_k^N with $k \in \mathcal{K}$ can be written as

$$\begin{cases} dz_k^N(t) = \Big[A_k z_k^N(t) + BR^{-1}B^T \bar{p}_k^N(t) - (BR^{-1}L - F)u_0(t)\Big]dt \\ \qquad\qquad + (1/N_k)\sum_{i\in\mathcal{N}_k} DdW_i(t) \\ d\bar{p}_k^N(t) = \Big[-A_k^T \bar{p}_k^N(t) + Q(z_k^N(t) - Hz(t))\Big]dt \\ \qquad\qquad + (1/N_k)\sum_{i\in\mathcal{N}_k} r_i(t)dW_i(t) \\ z_k^N(0) = (1/N_k)\sum_{i\in\mathcal{N}_k} x_i(0),\ \bar{p}_k^N(T) = 0, \end{cases}$$

where $\bar{p}_k^N(t) = \frac{1}{N_k}\sum_{i\in\mathcal{N}_k} p_i(t)$. Equivalently, the state-feedback representation for z_k^N can be written as

$$\begin{cases} dz_k^N(t) = \Big[\mathbb{G}_k(t) z_k^N(t) + BR^{-1}B^T \bar{\phi}_k^N(t) - (BR^{-1}L - F)u_0(t)\Big]dt \\ \qquad\qquad + (1/N_k)\sum_{i\in\mathcal{N}_k} DdW_i(t) \\ d\bar{\phi}_k^N(t) = \Big[-\mathbb{G}_k^T(t)\bar{\phi}_k^N(t) - QHz(t) - Z_k(t)(BR^{-1}L - F)u_0(t)\Big]dt \\ \qquad\qquad + (1/N_k)\sum_{i\in\mathcal{N}_k}(Z_i(t)D + r_i(t))dW_i(t) \\ z_k^N(0) = (1/N_k)\sum_{i\in\mathcal{N}_k} x_i(0), \quad \bar{\phi}_k^N(T) = 0, \end{cases}$$

where $\bar{\phi}_k^N(t) = \frac{1}{N_k}\sum_{i\in\mathcal{N}_k}\phi_i(t)$.

Note that the Brownian motions for $\mathcal{P}_i$, $1 \leq i \leq N$, are independent of each other, and are also independent of the initial conditions, and we have $\mathbb{E}\int_0^T \|r_i(t)\|^2 dt < \infty$. Moreover, under Assumption 5.1, $N \to \infty$ implies $N_k \to \infty$ for all $k \in \mathcal{K}$. Then from the strong law of large numbers (SLLN) [122], $(1/N_k)\sum_{i\in\mathcal{N}_k} DdW_i(t)$ and $(1/N_k)\sum_{i\in\mathcal{N}_k} r_i(t)dW_i(t)$ are negligible as $N \to \infty$ (as mentioned, that implies $N_k \to \infty$), in view of which $z_k(t) = \lim_{N_k\to\infty} z_k^N(t)$ and $\bar{p}_k(t) = \lim_{N_k\to\infty}\bar{p}_k^N(t)$ are well-defined processes. Therefore, since the mean field $x^N(t)$ was replaced by $z(t)$ in (5.9), we may denote the mean field limit $z(t) = \lim_{N\to\infty} x^N(t) = \lim_{N\to\infty}\sum_{k=1}^{K}\pi_k^N z_k^N(t) = \sum_{k=1}^{K}\pi_k z_k(t)$, which is a well-defined process.

Hence, z_k and $\bar{p}_k$ for $k \in \mathcal{K}$ can be written as

$$\begin{cases} dz_k(t) = \Big[A_k z_k(t) + BR^{-1}B^T\bar{p}_k(t) - (BR^{-1}L - F)u_0(t)\Big]dt \\ d\bar{p}_k(t) = \Big[-A_k^T\bar{p}_k(t) + Q(z_k(t) - Hz(t))\Big]dt \\ z_k(0) = \bar{x},\ \bar{p}_k(T) = 0. \end{cases} \tag{5.15}$$

It should be noted that the actual mean field stochastic process x^N is now captured by the stochastic process z with $z(t) = \sum_{k=1}^K \pi_k z_k(t)$ that depends on the leader's open-loop strategy $u_0 \in \mathcal{U}_0$ (note that u_0 is adapted to the filtration $\mathcal{F}_t$). By using the state-feedback representation given in (5.13), we may obtain the equivalent mean field process

$$\begin{cases} dz_k(t) = \Big[\mathbb{G}_k(t)z_k(t) + BR^{-1}B^T\bar{\phi}_k(t) - (BR^{-1}L - F)u_0(t)\Big]dt \\ d\bar{\phi}_k(t) = \Big[-\mathbb{G}_k^T(t)\bar{\phi}_k(t) - QHz(t) - Z_k(t)(BR^{-1}L - F)u_0(t)\Big]dt \\ z_k(0) = \bar{x},\ \bar{\phi}_k(T) = 0, \end{cases} \tag{5.16}$$

where $\bar{\phi}_k(t) = \lim_{N_k \to \infty} \bar{\phi}_k^N(t)$. Note that the state-feedback mean field representation is also dependent on the leader's strategy $u_0 \in \mathcal{U}_0$.

Proposition 5.1. *Given u_0, the stochastic mean field process (5.15) (and therefore (5.16)) admits a unique solution for all $k \in \mathcal{K}$ if there exists a solution of the following RDE*

$$-\frac{d\tilde{Z}(t)}{dt} = \mathbb{A}^T\tilde{Z}(t) + \tilde{Z}(t)\mathbb{A} + \mathbb{Q}_3 - \tilde{Z}(t)\mathbb{B}_2\tilde{Z}(t),\ \tilde{Z}(T) = 0, \tag{5.17}$$

where $\mathbb{A}$, $\mathbb{Q}_3$ and $\mathbb{B}_2$ are defined in Appendix D.1.

Proof. We can see that $\bar{p}_k$ is coupled with z_k and the mean field term $z(t) = \sum_{k=1}^K \pi_k z_k(t)$. Let $\tilde{z}$ and $\tilde{p}$ be the column vectors associated with z_k and $\bar{p}_k$, $k = 1, 2, ..., K$, respectively. Then $\tilde{z}$ and $\tilde{p}$ can be written as

$$d\tilde{z}(t) = \Big[\mathbb{A}\tilde{z}(t) + \mathbb{B}_2\tilde{p}(t) - \mathbb{BL}u_0(t)\Big]dt,\ d\tilde{p}(t) = \Big[-\mathbb{A}^T\tilde{p}(t) + \mathbb{Q}_3\tilde{z}(t)\Big]dt,$$

where $\mathbb{BL} = \mathbf{1}_K \otimes (BR^{-1}L - F)$. Note that $\tilde{z}(0) = \mathbf{1}_K \otimes \bar{x}$ and $\tilde{p}(T) = \mathbf{0}_{nK}$. It can be shown that $\tilde{p}(t) = -\tilde{Z}(t)\tilde{z}(t) + \tilde{\phi}(t)$, where

$$d\tilde{\phi}(t) = \Big[-(\mathbb{A}^T - \tilde{Z}(t)\mathbb{B}_2)\tilde{\phi}(t) - \tilde{Z}(t)\mathbb{BL}u_0\Big]dt,\ \tilde{\phi}(T) = 0.$$

Then from [121, Theorem 4.1], the existence and uniqueness of a solution to (5.15) (hence, to (5.16)) follow from the existence of a unique solution of (5.17). This completes the proof. □

The RDE given in (5.17) is not symmetric and positive semi-definite, since $\mathbb{Q}_3$ is neither symmetric nor positive semi-definite. Therefore, in general, it may not admit a unique solution. We do not provide conditions for the existence of solution of (5.17), since this issue will be studied together with the leader's problem in Section 5.4, where conditions that guarantee existence and uniqueness of the leader's optimal solution as well as the solution of the mean field process are provided in terms of a linear matrix inequality. Therefore, in the rest of this section, we assume that the solution of the mean field process in (5.15) (or (5.16)) exists.

The following result shows that (5.15) or (5.16) is indeed the best estimate of the actual mass behavior when N is arbitrarily large.

Proposition 5.2. *Suppose that Assumption 5.1 holds. Then for any* $u_0 \in \mathcal{U}_0$*, we have*

$$\mathbb{E}\int_0^T \|x^N(t) - z(t)\|^2 dt = O\Big(\frac{1}{N} + \epsilon_N^2\Big),$$

where $\epsilon_N = \sup_{k\in\mathcal{K}} |\pi_k^N - \pi_k|$*. Moreover, we have*

$$\sup_{0\leq t\leq T} \mathbb{E}[\|x^N(t) - z(t)\|^2] = O(\frac{1}{N} + \epsilon_N^2).$$

Proof. We prove the first statement only, since the second statement can be shown in a similar way. We use the state-feedback representation given in (5.16). First, note that

$$\begin{aligned}\|x^N(t) - z(t)\|^2 &= \Big\|\sum_{k=1}^K \pi_k^N z_k^N(t) - \sum_{k=1}^K \pi_k z_k(t)\Big\|^2 \\ &\leq 2\Big\|\sum_{k=1}^K \pi_k^N (z_k^N(t) - z_k(t))\Big\|^2 + 2\epsilon_N^2\Big\|\sum_{k=1}^K z_k(t)\Big\|^2.\end{aligned}$$

Let $e_k(t) = z_k^N(t) - z_k(t)$ and $\bar{e}_k(t) = \bar{\phi}_k^N(t) - \bar{\phi}_k(t)$, where $\bar{e}_k(T) = 0$ for all $k \in \mathcal{K}$. Moreover, we have $\mathbb{E}[e_k(0)] = 0$ and $\mathbb{E}[\|e_k(0)\|^2] < \infty$ for all $k \in \mathcal{K}$.

Then

$$\mathbb{E}\int_0^T \|x^N(t)-z(t)\|^2 dt$$
$$\leq 2\mathbb{E}\int_0^T \Big\|\sum_{k=1}^K \pi_k^N e_k(t)\Big\|^2 dt + 2\epsilon_N^2 \mathbb{E}\int_0^T \sum_{k=1}^K \|z_k(t)\|^2 dt, \tag{5.18}$$

where we make use of the fact that Brownian motions of the followers are independent of each other to get the second term in the RHS of the inequality. Then, for the second term, in view of the fact that $\mathbb{E}\int_0^T \|z_k(t)\|^2 dt < \infty$ for all $k \in \mathcal{K}$, we can show that $\epsilon_N^2 \mathbb{E}\int_0^T \sum_{k=1}^K \|z_k(t)\|^2 dt = O(\epsilon_N^2)$ as $N \to \infty$.

Now, for the first term of (5.18), we have

$$de_k(t) = [\mathbb{G}_k(t)e_k(t) + BR^{-1}B^T\bar{e}_k(t)]dt + (1/N_k)\sum_{i\in\mathcal{N}_k} DdW_i(t)$$
$$d\bar{e}_k(t) = -\mathbb{G}_k^T(t)\bar{e}_k(t)dt + (1/N_k)\sum_{i\in\mathcal{N}_k}(Z_i(t)D + r_i(t))dW_i(t).$$

Let $\Phi_i(t,s)$ be the state transition matrix associated with $\mathbb{G}_i(t)$, that is, $d\Phi_i(t,s) = \mathbb{G}_i(t)\Phi_i(t,s)dt$. We have the following facts from linear system theory:

$$d\Phi_i^T(s,t) = -\mathbb{G}_i^T(t)\Phi_i^T(s,t)dt,\ d\Phi_i^T(t,s) = \Phi_i^T(t,s)\mathbb{G}_i^T(t)dt$$
$$\Phi_i^T(0,t)\Phi_i^T(s,0) = (\Phi_i(s,0)\Phi_i(0,t))^T = \Phi_i^T(s,t),$$

where $0 \leq s,t \leq T$. For any $k \in \mathcal{K}$, by using Itô formula, we have

$$d\Phi_k^T(t,0)\bar{e}_k(t) = \Phi_k^T(t,0)(1/N_k)\sum_{i\in\mathcal{N}_k}(Z_i(t)D + r_i(t))dW_i(t).$$

Then,

$$-\Phi_k^T(t,0)\bar{e}_k(t) = \int_t^T \Phi_k^T(\tau,0)(1/N_k)\sum_{i\in\mathcal{N}_k}(Z_i(t)D + r_i(t))dW_i(t),$$

which implies

$$\bar{e}_k(t) = -\int_t^T \Phi_k^T(\tau,t)(1/N_k)\sum_{i\in\mathcal{N}_k}(Z_i(t)D + r_i(t))dW_i(t).$$

Since $\bar{e}_k(t)$ has to be an adapted solution with respect to $\mathcal{G}_t^k = \vee_{i\in\mathcal{N}_k}\mathcal{F}_t^i$, we must have $e_k(t) = \mathbb{E}[e_k(t)|\mathcal{G}_t^k]$ for all $k \in \mathcal{K}$, which implies that $\bar{e}_k(t) \equiv 0$ for all $k \in \mathcal{K}$. Due to the fact that $\sup_{k\in\mathcal{K},0\leq t\leq T}\|Z_k(t)\| \leq C$ and $\Phi_k(t,\tau)$ is continuous in t and τ, there exists a constant $C_1 > 0$ independent of N such that $\sup_{k\in\mathcal{K},0\leq t,\tau\leq T}\|\Phi_k(t,\tau)\| \leq C_1 < \infty$ [85, Remark A.4]. Hence, $\mathbb{E}\int_0^T \|e_k(t)\|^2 dt \leq C_2 < \infty$ for all $k \in \mathcal{K}$. Now, due to the fact that $\lim_{N\to\infty}\pi_k^N = \pi_k > 0$ from Assumption 5.1, it can be shown that

$$\frac{1}{N^2}\mathbb{E}\int_0^T \Big\|\sum_{k=1}^K N_k e_k(t)\Big\|^2 dt = \frac{1}{N^2}\int_0^T \sum_{i=1}^N \mathbb{E}[\|e_i(t)\|^2]dt \leq \frac{C_2 N}{N^2},$$

where the equality follows from the independence of the Brownian motions. Then we have the desired result. This completes the proof. □

5.3.3 Optimality of N followers: The ϵ-Nash Equilibrium

This subsection shows that the set of N optimal decentralized strategies for $\mathcal{P}_i$, $1 \leq i \leq N$ obtained in Section 5.3.2, constitutes an ϵ-Nash equilibrium for $\mathcal{P}_i$, $1 \leq i \leq N$, given $u_0 \in \mathcal{U}_0$. The definition of an ϵ-Nash equilibrium is given as follows.

Definition 5.1. *The set of strategies, $\{\bar{u}_i \in \mathcal{U}_i(u_0),\ 1 \leq i \leq N\}$, constitutes an ϵ-Nash equilibrium with respect to $\{J_i^N,\ 1 \leq i \leq N\}$, if there exists $\epsilon \geq 0$ such that $J_i^N(\bar{u}_i, \bar{u}_{-i}, u_0) \leq \inf_{u_i\in\mathcal{U}_i(u_0)} J_i^N(u_i, \bar{u}_{-i}, u_0) + \epsilon$, for all i, $1 \leq i \leq N$.*

We now recall the dynamics and the corresponding optimal decentralized controller for the follower $\mathcal{P}_i$, $1 \leq i \leq N$, given in Lemma 5.1:

$$\begin{cases} dx_i^*(t) = \Big[A_i x_i^*(t) + BR^{-1}B^T p_i(t) - (BR^{-1}L - F)u_0(t)\Big]dt + DdW_i(t) \\ dp_i(t) = \Big[-A_i^T p_i(t) + Q(x_i^*(t) - Hz(t))\Big]dt + r_i(t)dW_i(t) \\ x_i^*(0) = x_i(0),\ p_i(T) = 0 \\ u_i^*(t) = R^{-1}B^T p_i(t) - R^{-1}Lu_0(t), \end{cases} \tag{5.19}$$

where p_i now depends on the mean field process $z(t) = \sum_{k=1}^K z_k(t)$ determined in (5.15). As already mentioned in Remark 5.1, the above optimal controller is decentralized in terms of the local information and the arbitrary strategy of the leader. It should be noted that in Definition 5.1, the set of the follower's (say i'th) admissible strategies, $\mathcal{U}_i(u_0)$, was not restricted to be decentralized.

Theorem 5.1. *Suppose that Assumption 5.1 holds. For any $u_0 \in \mathcal{U}_0$, the set of N strategies in (5.19) (or (5.14)), i.e., $u^{N*} = \{u_i^*, 1 \leq i \leq N\}$, constitutes an ϵ-Nash equilibrium for $\{\mathcal{P}_i,\ 1 \leq i \leq N\}$, that is, for any i, $1 \leq i \leq N$, we have*

$$J_i(u_i^*, u_{-i}^*, u_0) \leq \inf_{u_i \in \mathcal{U}_i(u_0)} J_i(u_i, u_{-i}^*, u_0) + \epsilon,$$

where $\epsilon = O(\frac{1}{\sqrt{N}} + \epsilon_N)$ and ϵ_N is defined in Proposition 5.2.

Proof. The proof consists of two parts. First, by using Cauchy-Schwarz inequality, we have

$$\begin{aligned} J_i^N(u_i^*, u_{-i}^*, u_0) &\leq \bar{J}_i(u_i^*) + \epsilon^2(N) + 2\Big(\mathbb{E}\int_0^T \|x_i^*(t) - z(t)\|_Q^2 dt\Big)^{\frac{1}{2}} \epsilon(N) \\ &= \bar{J}_i(u_i^*) + O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big), \end{aligned} \tag{5.20}$$

where $\epsilon(N) := (\mathbb{E}\int_0^T \|x^N(t) - z(t)\|^2 dt)^{1/2} = O(\frac{1}{\sqrt{N}} + \epsilon_N)$ in view of Proposition 5.2, and the equality follows from Proposition 5.2 and the fact that $x_i^* \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)$. Note that the above condition implies that for any $u_0 \in \mathcal{U}_0$ and i,

$$\Big| J_i^N(u_i^*, u_{-i}^*, u_0) - \bar{J}_i(u_i^*) \Big| = O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big).$$

On the other hand, since $\mathbb{E}\int_0^T \|x_i^*(t)\|^2 dt < \infty$ and $\mathbb{E}\int_0^T \|z(t)\|^2 dt < \infty$, we can show that $J_i^N(u_i^*, u_{-i}^*, u_0) \leq C < \infty$, where C is independent of N. Since we have $\inf_{u_i \in \mathcal{U}_i(u_0)} J_i^N(u_i, u_{-i}^*, u_0) \leq J_i^N(u_i^*, u_{-i}^*, u_0)$, we may consider the controller $u_i \in \mathcal{U}_i(u_0)$ that satisfies $\mathbb{E}\int_0^T \|x_i(t)\|^2 dt < \infty$ for all i. Then we can show that

$$J_i^N(u_i, u_{-i}^*, u_0) \geq \bar{J}_i(u_i) + I_1 + I_2,$$

where

$$
\begin{aligned}
I_1 &= 2\mathbb{E}\int_0^T (x_i(t) - Hz(t))^T QH(z(t) - x^N(t))dt \\
I_2 &= \frac{2}{N}\mathbb{E}\int_0^T (x_i(t) - Hz(t))^T QH(x_i^*(t) - x_i(t))dt.
\end{aligned}
$$

By using Cauchy-Schwarz inequality, Proposition 5.1, and the fact that $\mathbb{E}\int_0^T \|x_i(t)\|^2 dt < \infty$ and $\mathbb{E}\int_0^T \|z(t)\|^2 dt < \infty$ for all i, we can show that

$$
\begin{aligned}
\bar{J}_i(u_i^*) &\leq \bar{J}_i(u_i) \\
&\leq J_i^N(u_i, u_{-i}^*, u_0) + |I_1| + |I_2| \\
&\leq J_i^N(u_i, u_{-i}^*, u_0) + O\Big(\frac{1}{N}\Big) \\
&\quad + 2\sqrt{\|H\|\|Q\|}\Big(\mathbb{E}\int_0^T \|x_i(t) - Hz(t)\|_Q^2 dt\Big)^{1/2}\epsilon(N) \\
&= J_i^N(u_i, u_{-i}^*, u_0) + O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big),
\end{aligned} \tag{5.21}
$$

where $\epsilon(N) := (\mathbb{E}\int_0^T \|x^N(t) - z(t)\|^2 dt)^{1/2} = O(\frac{1}{\sqrt{N}} + \epsilon_N)$ and the first inequality follows from the fact that u_i^* is the corresponding minimizing solution of the follower's local optimal control problem. Then, from (5.20) and (5.21), we have the desired result. This completes the proof. □

We now discuss the relationship between the mean field approach in this section and the standard approach discussed in Section 5.2.2. For any $u_0 \in \mathcal{U}_0$ and as $N \to \infty$, in view of the FBSDE given in (5.6), we can show that for all $k \in \mathcal{K}$

$$
\bar{p}_k = \lim_{N_k \to \infty} (1/N_k) \sum_{i \in \mathcal{N}_k} s_i^N(t),\ z_k(t) = \lim_{N_k \to \infty} (1/N_k) \sum_{i \in \mathcal{N}_k} x_i'(t),
$$

since $(1/N)H^T Q x_i'(t) = o(1/N)$ and $(1/N)H^T QHx^N(t) = o(1/N)$ in (5.6) are negligible as $N \to \infty$. This implies that the mean field process obtained by the standard approach is identical to that in (5.15). Moreover, we have $\lim_{N\to\infty} s_i^N = p_i$ and $\lim_{N\to\infty} x_i' = x_i^*$ for all i, where x_i^* and p_i are obtained in Lemma 5.1. This shows that when $N \to \infty$, the ϵ-Nash equilibrium in Theorem 5.1 is the same as the centralized Nash equilibrium in Section 5.2.2 (that is the set of the centralized strategies u_i', $1 \leq i \leq N$, in (5.5)), since

$u_i^* = u_i'$ for all i, provided that the solution of the corresponding FBSDE in (5.6) exists for all N.

5.4 Leader's Problem

In this section, we solve the leader's problem to obtain an approximated Stackelberg equilibrium.

5.4.1 Leader's Local Optimal Control Problem

Due to the nature of the Stackelberg game under consideration, the leader's local optimal control problem is to minimize

$$\bar{J}_0(u_0) = \mathbb{E}\int_0^T \Big[\|x_0(t) - H_0 z(t)\|^2_{Q_0} + \|u_0(t)\|^2_{R_0}\Big] dt \tag{5.22}$$

subject to

$$dx_0(t) = [A_0 x_0(t) + B_0 u_0(t)]dt + D_0 dW_0(t), \tag{5.23}$$

and

$$\begin{cases} dz_k(t) = \Big[A_k z_k(t) + BR^{-1}B^T \bar{p}_k(t) - (BR^{-1}L - F)u_0(t)\Big] dt \\ d\bar{p}_k(t) = \Big[-A_k^T \bar{p}_k(t) + Q(z_k(t) - Hz(t))\Big] dt \\ z_k(0) = \bar{x},\ \bar{p}_k(T) = 0,\ k = 1, 2, ..., K, \end{cases} \tag{5.24}$$

where $\mathbb{E}[x_0(0)] = \bar{x}_0$ and $\mathbb{E}[\|x_0(0)\|^2] < \infty$. In (5.22), the mean field term is replaced with the approximated mean field behavior of the followers, $z(t) = \sum_{k=1}^K \pi_k z_k(t)$, which is dependent on the leader's arbitrary strategy $u_0 \in \mathcal{U}_0$ as can be seen from (5.24). Note that the leader's local optimal control problem has $2K$ additional constraints induced by the mean field approximation in Section 5.3, which is independent of N. Hence, as expected, the optimization problem (5.22) is much more tractable than the original optimal control problem faced by $\mathcal{P}_0$ in (5.7), although the optimal control problem in (5.22) is still not standard, since it has additional constraints that have initial and boundary conditions.

Note that constraints in (5.24) can be replaced by the state-feedback representation form, which is given below for convenience:

$$\begin{cases} dz_k(t) = \Big[\mathbb{G}_k(t)z_k(t) + BR^{-1}B^T\bar{\phi}_k(t) - (BR^{-1}L - F)u_0(t)\Big]dt \\ d\bar{\phi}_k(t) = \Big[-\mathbb{G}_k^T(t)\bar{\phi}_k(t) - QHz(t) - Z_k(t)(BR^{-1}L - F)u_0(t)\Big]dt \\ z_k(0) = \bar{x},\ \bar{\phi}_k(T) = 0,\ k = 1, 2, ..., K. \end{cases} \tag{5.25}$$

The following lemma solves the nonstandard optimal control problem in (5.22); the proof is given in Appendix D.3.

Lemma 5.2. *For the local optimal control problem for $\mathcal{P}_0$ in (5.22), there exists a unique optimal controller $u_0^* \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^q)$. Moreover, $(x_0^*, u_0^*) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n+p})$ is the corresponding optimal solution if and only if*

$$u_0^*(t) = R_0^{-1}B_0^T p_0(t) - R_0^{-1}(L^T R^{-1}B^T - F^T)\sum_{k=1}^{K}\pi_k\lambda_k(t), \tag{5.26}$$

where

$$\begin{cases} dx_0^*(t) = \Big[A_0x_0^*(t) + B_0R_0^{-1}B_0^Tp_0(t) \\ \qquad\qquad - B_0R_0^{-1}(L^TR^{-1}B^T - F^T)\sum_{k=1}^{K}\pi_k\lambda_k(t)\Big]dt + D_0dW_0(t) \\ dp_0(t) = \Big[-A_0^Tp_0(t) + Q_0(x_0^*(t) - H_0z(t))\Big]dt + q_0(t)dW_0(t) \\ d\lambda_k(t) = \Big[-A_k^T\lambda_k(t) + H_0^TQ_0(H_0z(t) - x_0^*(t)) - Q\xi_k(t) \\ \qquad\qquad + H^TQ\sum_{k=1}^{K}\pi_k\xi_k(t)\Big]dt + q_k(t)dW_0(t) \\ d\xi_k(t) = \Big[A_k\xi_k(t) - BR^{-1}B^T\lambda_k(t)\Big]dt \\ dz_k(t) = \Big[A_kz_k(t) + BR^{-1}B^T\bar{p}_k(t) - (BR^{-1}L - F)R_0^{-1}B_0^Tp_0(t) \\ \qquad\qquad + (BR^{-1}L - F)R_0^{-1}(L^TR^{-1}B^T - F^T)\sum_{k=1}^{K}\pi_k\lambda_k(t)\Big]dt \\ d\bar{p}_k(t) = \Big[-A_k^T\bar{p}_k(t) + Q(z_k(t) - Hz(t))\Big]dt \\ x_0^*(0) = x_0(0),\ \xi_k(0) = 0,\ z_k(0) = \bar{x} \\ p_0(T) = 0,\ \lambda_k(T) = 0,\ \bar{p}_k(T) = 0,\ k = 1, 2, ..., K. \end{cases} \tag{5.27}$$

Remark 5.2. *In (5.27), ξ_k and λ_k, $k = 1, 2, ..., K$, are additional $\mathcal{F}_t$ adapted forward-backward processes (or adjoint processes) generated by additional constraints z_k and $\bar{p}_k$, $k = 1, 2, ..., K$.*

Bearing in mind that in the follower's problem, the corresponding optimal solution exists if and only if the mean field limit z and the leader's optimal control u_0^* exist. In view of Lemma 5.2, we need to identify conditions for which there exists a set of unique solutions of the FBSDEs in (5.27).

Let

$$\mathcal{X}(t) = \begin{pmatrix} x_0^{*T}(t) & \xi_1^T(t) & \dots & \xi_K^T(t) & z_1^T(t) & \dots & z_K^T(t) \end{pmatrix}^T$$
$$\mathcal{Y}(t) = \begin{pmatrix} p_0^T(t) & \lambda_1^T(t) & \dots & \lambda_K^T(t) & \bar{p}_1^T(t) & \dots & \bar{p}_K^T(t) \end{pmatrix}^T,$$

where $\mathcal{X}(0) = (x_0^T(0), \mathbf{0}_{nK}^T, \mathbf{1}_K^T \otimes \bar{x}^T)^T$ and $\mathcal{Y}(T) = \mathbf{0}_{2nK+n}$. Note that $\mathcal{X}$ and $\mathcal{Y}$ are $(2nK+n)$-dimensional vector-valued forward and backward processes, respectively. Define

$$\mathcal{A}_1 = \text{diag}\{A_0, \mathbb{A}, \mathbb{A}\},\ \mathcal{B}_1 = \begin{pmatrix} B_0 R_0^{-1} B_0^T & \mathbb{B}_1 & \mathbf{0}_{n\times Kn} \\ \mathbf{0}_{nK\times n} & -\mathbb{B}_2 & \mathbf{0}_{nK\times Kn} \\ \mathbb{B}_3 & \mathbb{B}_K & \mathbb{B}_2 \end{pmatrix},\ \mathcal{D}_1 = \begin{pmatrix} D_0 \\ \mathbf{0}_{nK\times q} \\ \mathbf{0}_{nK\times q} \end{pmatrix}$$
$$\mathcal{A}_2 = \begin{pmatrix} Q_0 & \mathbf{0}_{n\times nK} & \mathbb{Q}_1 \\ \mathbb{Q}_2 & -\mathbb{Q}_3^T & \mathbb{Q}_K \\ \mathbf{0}_{nK\times n} & \mathbf{0}_{nK\times nK} & \mathbb{Q}_3 \end{pmatrix},\ \mathcal{B}_2 = -\mathcal{A}_1^T,\ \mathcal{D}_2 = \begin{pmatrix} q_0(t) \\ \bar{q}(t) \\ \mathbf{0}_{nK\times q} \end{pmatrix},$$

where the corresponding block matrices are defined in Appendix D.1. We can show that

$$d\mathcal{X}(t) = [\mathcal{A}_1\mathcal{X} + \mathcal{B}_1\mathcal{Y}]dt + \mathcal{D}_1 dW_0(t)$$
$$d\mathcal{Y}(t) = [\mathcal{A}_2\mathcal{X} + \mathcal{B}_2\mathcal{Y}]dt + \mathcal{D}_2 dW_0(t).$$

It can be shown that $\mathcal{X}$ and $\mathcal{Y}$ satisfy the following affine transformation

$$\mathcal{Y}(t) = -\Lambda(t)\mathcal{X}(t) + \mathcal{V}(t), \tag{5.28}$$

where $\Lambda(t)$ is a solution to the following RDE

$$-\frac{d\Lambda(t)}{dt} = \Lambda(t)\mathcal{A}_1 - \mathcal{B}_2\Lambda(t) + \mathcal{A}_2 - \Lambda(t)\mathcal{B}_1\Lambda(t),\ \Lambda(T) = 0, \tag{5.29}$$

and $\mathcal{V}(t)$ satisfies $\mathcal{V}(T) = 0$ and

$$d\mathcal{V}(t) = [\mathcal{B}_2 + \Lambda(t)\mathcal{B}_1]\mathcal{V}(t)dt + [\mathcal{D}_2 + \Lambda(t)\mathcal{D}_1]dW_0(t).$$

Note that $\mathcal{V}$ is decoupled from $\mathcal{X}$. It is easy to see that the RDE in (5.29) is *nonsymmetric*; hence in general, it may not admit a solution [121, 123]. Since $\mathcal{X}$ and $\mathcal{Y}$ satisfy the affine transformation (5.28), $\mathcal{X}$ and $\mathcal{Y}$ admit a unique solution if the nonsymmetric RDE in (5.29) has a unique solution for all $t \in [0, T]$ [121, Theorem 4.1] (or see [92]). We now identify conditions for the existence of a unique solution of (5.29).

The first result is given below, which follows from direct computation, or see [121, Theorem 4.3].

Proposition 5.3. *Suppose that we have*

$$\det\left\{\begin{pmatrix} I & \boldsymbol{0} \end{pmatrix} e^{\bar{\mathcal{A}}(t-T)} \begin{pmatrix} I \\ \boldsymbol{0} \end{pmatrix}\right\} > 0,\ \forall t \in [0, T],\ \bar{\mathcal{A}} = \begin{pmatrix} \mathcal{A}_1 & -\mathcal{B}_1 \\ -\mathcal{A}_2 & \mathcal{B}_2 \end{pmatrix}. \quad (5.30)$$

Then the RDE in (5.29) admits a unique solution $\Lambda(t)$ *for all* $t \in [0, T]$ *with* $\Lambda(T) = 0$*, which can be written as*

$$\Lambda(t) = \left[\begin{pmatrix} \boldsymbol{0} & I \end{pmatrix} e^{\bar{\mathcal{A}}(t-T)} \begin{pmatrix} I \\ \boldsymbol{0} \end{pmatrix}\right]\left[\begin{pmatrix} I & \boldsymbol{0} \end{pmatrix} e^{\bar{\mathcal{A}}(t-T)} \begin{pmatrix} I \\ \boldsymbol{0} \end{pmatrix}\right]^{-1} = \Pi(t)\Psi^{-1}(t),$$

where $\Pi(t)$ *and* $\Psi(t)$ *are defined in Lemma D.1 in Appendix D.2.*

Note that the condition in (5.30) requires computation of the matrix exponential function ($(4nK+2n) \times (4nK+2n)$-dimensional) and its determinant for all t; therefore, it is sometimes hard to check. The following result provides an easy-to-check condition for existence and uniqueness of the solution of (5.29), which can be viewed as a modified version of [123, Theorem 3.11].

Proposition 5.4. *Let*

$$\Upsilon(W_1, W_2) = \begin{pmatrix} W_1\mathcal{A}_1 - W_2\mathcal{A}_2 & -W_1\mathcal{B}_1 + \mathcal{A}_1^T W_2 + W_2\mathcal{B}_2 - \mathcal{A}_2^T W_2 \\ 0 & -\mathcal{B}_1^T W_2 + \mathcal{B}_2^T W_2 \end{pmatrix},$$

where $W_1 = W_1^T > 0$ *and* $W_2 = W_2^T$*. Suppose* $\Upsilon + \Upsilon^T < 0$*. Then, the RDE in (5.29) admits a unique solution* $\Lambda(t)$ *for all* $t \in [0, T]$ *with* $\Lambda(T) = 0$*.*

Proof. The local existence and uniqueness theorem of differential equations implies that there exists a solution to the differential equation in Lemma D.1 in Appendix D.2 for t sufficiently close to T, that is, $\Psi(t)$ is invertible for t sufficiently close to T. Therefore, it suffices to show that if $\Upsilon + \Upsilon^T < 0$ holds with $W_1 = W_1^T > 0$ and $W_2 = W_2^T$, $\Psi(t)$ is invertible for all $t \in [0, T]$.

For any $z \in \mathbb{R}^{2nK+n}$ and $z \neq 0$, define

$$\begin{aligned} V(t,z) = \; & z^T(\Psi^T(t)W_1\Psi(t) + \Psi^T(t)W_2\Pi(t) \\ & + \Pi^T(t)W_2\Psi(t) + \Pi^T(t)W_2\Pi(t))z. \end{aligned}$$

Clearly, the matrix-valued function in $V(t,z)$ is symmetric, that is, $V(t,z) = V^T(t,z)$, and in view of the terminal condition, we have $V(T,z) = z^T W_1 z > 0$. Let $\bar{z} = (z^T\Psi^T(t), z^T\Pi^T(t))^T$. It can be shown that ($t$ argument is suppressed)

$$\begin{aligned} \frac{dV}{dt} &= 2z^T\Psi^T W_1 \frac{d\Psi}{dt} z + 2z^T \frac{d\Psi^T}{dt} W_2 \Pi z + 2z^T \Psi W_2 \frac{d\Pi}{dt} z + 2z^T \frac{d\Pi^T}{dt} W_2 \Pi z \\ &= 2z^T\Psi^T W_1(\mathcal{A}_1\Psi - \mathcal{B}_1\Pi)z + 2z^T(\Psi^T\mathcal{A}_1^T - \Pi^T\mathcal{B}_1^T)W_2\Pi z \\ &\quad + 2z^T\Psi^T W_2(-\mathcal{A}_2\Psi + \mathcal{B}_2\Pi)z + 2z^T(-\Psi^T\mathcal{A}_2^T + \Pi^T\mathcal{B}_2^T)W_2\Pi z \\ &= \bar{z}^T(\Upsilon + \Upsilon^T)\bar{z} < 0. \end{aligned}$$

This implies that $V(t,z)$ is monotonically decreasing from $(-\infty, T]$, and since $V(T,z) = z^T W_1 z > 0$, $V(t,z) > 0$ for all $t \in [0,T]$. This also implies that the matrix-valued function in $V(t,z)$ is positive definite for all $t \in [0,T]$; therefore, $\Psi(t)$ is invertible for all $t \in [0,T]$. This completes the proof. □

Remark 5.3. *Let* $\mathcal{L}(W_1, W_2) = diag\{\Upsilon + \Upsilon^T, -W_1\}$, *and* $\mathcal{C} = \{W_1 = W_1^T > 0, W_2 = W_2^T : \mathcal{L}(W_1, W_2) < 0\}$. *We can easily check that* $\mathcal{L}(W_1, W_2) < 0$ *is a linear matrix inequality (LMI), and* $\mathcal{C}$ *is convex. Therefore, the condition in Proposition 5.4 can be checked by using standard semidefinite programming by identifying feasibility of the corresponding LMI [115].*

We write $\Lambda(t)$ in the partition form:

$$\Lambda(t) = \begin{pmatrix} \Lambda_{11}(t) & \Lambda_{12}(t) & \Lambda_{13}(t) \\ \Lambda_{21}(t) & \Lambda_{22}(t) & \Lambda_{23}(t) \\ \Lambda_{31}(t) & \Lambda_{32}(t) & \Lambda_{33}(t) \end{pmatrix},$$

where Λ_{11} is $n \times n$, and Λ_{22} and Λ_{33} are $nK \times nK$ matrices. Then it is easy to check that the existence of a solution $\Lambda(t)$ (especially Λ_{33}) implies the existence of a solution of the RDE in (5.17), which guarantees the existence of a solution of the mean field process in Proposition 5.1.

Before concluding this subsection, we should note that under the optimal decentralized controller of the leader in (5.26), the approximated mean field process, z_k and $z(t) = \sum_{k=1}^{K} \pi_k z_k(t)$, can be obtained by simply taking the conditional expectation with respect to $\mathcal{F}_t^0$, i.e., $z_k(t) = \mathbb{E}[z_k^N(t)|\mathcal{F}_t^0]$, which is indeed equivalent to applying the SLLN in Section 5.3.2. This implies that under the optimal strategy of the leader, the approximated mean field is a stochastic process adapted to the leader's local information, which is similar to the major and minor problem considered in [84] and [85], but their stochastic mean field processes are *indirectly* affected by the major's strategy, since the control input of the major player does not appear in the corresponding mean field process.

5.4.2 Optimality of the Leader: The (ϵ_1, ϵ_2)-Stackelberg Equilibrium

This subsection shows that if the leader announces u_0^* obtained in Section 5.4.1 to the N followers, the set of the optimal decentralized strategies for the leader and the followers constitutes an approximated Stackelberg equilibrium. The definition of an (ϵ_1, ϵ_2)-Stackelberg equilibrium is as follows, which can be viewed as a modified version of the definition of an ϵ-Stackelberg equilibrium given in [61].

Definition 5.2. *Let $\bar{u}^N(\bar{u}_0) = \{\bar{u}_1(\bar{u}_0), ..., \bar{u}_N(\bar{u}_0),\ \bar{u}_i \in \mathcal{U}_i(\bar{u}_0),\ 1 \leq i \leq N\}$ and $\bar{u}_0 \in \mathcal{U}_0$. Then $\bar{u} = \{\bar{u}_0, \bar{u}^N(\bar{u}_0)\}$ is an (ϵ_1, ϵ_2)-Stackelberg equilibrium with respect to $\{J_i^N, 0 \leq i \leq N\}$ if the following two properties hold:*

(i) $\bar{u}^N(\bar{u}_0)$ constitutes an ϵ_1-Nash equilibrium under $\bar{u}_0$.

(ii) There exists $\epsilon_2 \geq 0$ such that

$$J_0^N(\bar{u}_0, \bar{u}^N(\bar{u}_0)) \leq \inf_{u_0 \in \mathcal{U}_0} J_0^N(u_0, \bar{u}^N(u_0)) + \epsilon_2.$$

Since, in Definition 5.2, the followers are ϵ-Nash followers, the above definition can also be viewed as an approximated version of the definition of the

Stackelberg equilibrium given in [89, 118].

We now recall the dynamics and the corresponding optimal decentralized controller for $\mathcal{P}_0$:

$$\begin{cases} dx_0^*(t) = \Big[A_0 x_0^*(t) + B_0 R_0^{-1} B_0^T p_0(t) \\ \qquad\qquad - B_0 R_0^{-1}(L^T R^{-1} B^T - F^T) \sum_{k=1}^K \pi_k \lambda_k(t)\Big] dt + D_0 dW_0(t) \\ dp_0(t) = \Big[-A_0^T p_0(t) + Q_0(x_0^*(t) - H_0 z(t))\Big] dt + q_0(t) dW_0(t) \\ u_0^*(t) = R_0^{-1} B_0^T p_0(t) - R_0^{-1}(L^T R^{-1} B^T - F^T) \sum_{k=1}^K \pi_k \lambda_k(t), \end{cases} \tag{5.31}$$

where the corresponding FBSDEs are given in Lemma 5.2. Let

$$u^* = \{u_0^*, u^{N*}(u_0^*)\},\ u^{N*}(u_0^*) = \{u_1^*(u_0^*), ..., u_N^*(u_0^*)\}, \tag{5.32}$$

where u_0^* is determined from (5.31) and $u^{N*}(u_0^*)$ is the set of the followers' optimal decentralized strategies in Theorem 5.1 when u_0 is replaced by u_0^*. Note that $u^{N*}(u_0^*)$ is the set of followers' optimal decentralized strategies when the leader announces his optimal decentralized strategy u_0^* to the N followers.

Theorem 5.2. *Suppose that Assumption 5.1 holds and the RDE in (5.29) has a unique solution $\Lambda(t)$ for all $t \in [0,T]$ with $\Lambda(T) = 0$. Then u^* given in (5.32) constitutes an (ϵ_1, ϵ_2)-Stackelberg equilibrium, where $\epsilon_1 = \epsilon_2 = O(\frac{1}{\sqrt{N}} + \epsilon_N)$ and ϵ_N is defined in Proposition 5.2.*

Proof. We first note that under u_0^*, $u^{N*}(u_0^*)$ constitutes an ϵ_1-Nash equilibrium where $\epsilon_1 = O(\frac{1}{\sqrt{N}} + \epsilon_N)$ in view of Theorem 5.1. Therefore, it suffices to show that the optimal decentralized strategy of the leader (5.31) with $u^{N*}(u_0^*)$ satisfies the second property in Definition 5.2. Let z_{-0} be the mean field process when the leader takes an arbitrary strategy $u_0 \in \mathcal{U}_0$. Let x_{-0}^N be the actual mass behavior when the followers are under $u^{N*}(u_0)$. Note that z_{-0} and x_{-0}^N are equivalent to those in Proposition 5.2, since Proposition 5.2 holds for any arbitrary strategy of the leader.

Similar to Theorem 5.1, we can show that

$$\Big| J_0^N(u_0^*, u^{N*}(u_0^*)) - \bar{J}_0(u_0^*) \Big| = O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big),$$

which implies

$$J_0^N(u_0^*, u^{N*}(u_0^*)) \leq \bar{J}_0(u_0^*) + O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big), \tag{5.33}$$

where $\bar{J}_0(u_0^*)$ is the minimum cost of the leader's local optimal control problem in Lemma 5.2. Note that $\bar{J}_0(u_0^*)$ is related to the mean field process $z(t)$ given in Lemma 5.2.

By virtue of Theorem 5.1 and Lemma 5.2, we have that $\mathbb{E}\int_0^T \|x_0^*(t)\|^2 dt < \infty$ and $\mathbb{E}\int_0^T \|x_i^*(t)\|^2 dt < \infty$ for all i; hence there exists a constant C, independent of N, such that $J_0^N(u_0^*, u^{N*}(u_0^*)) \leq C < \infty$. Since for any $u_0 \in \mathcal{U}_0$, we have $\inf_{u_0 \in \mathcal{U}_0} J_0^N(u_0, u^{N*}(u_0)) \leq J_0^N(u_0^*, u^{N*}(u_0^*))$, it suffices to consider $u_0 \in \mathcal{U}_0$ with the property that $\mathbb{E}\int_0^T \|x_0(t)\|^2 dt < \infty$.

Now, we have

$$J_0^N(u_0, u^{N*}(u_0)) \geq \bar{J}_0(u_0) + I,$$

where

$$I = 2\mathbb{E}\int_0^T (x_0(t) - z_{-0}(t))^T Q(z_{-0}(t) - H_0 x_{-0}^N(t))dt.$$

Then by using Cauchy-Schwarz inequality and Proposition 5.2, we can show that

$$\bar{J}_0(u_0^*) \leq \bar{J}_0(u_0) \leq J_0^N(u_0, u^{N*}(u_0)) + O\Big(\frac{1}{\sqrt{N}} + \epsilon_N\Big), \tag{5.34}$$

where the first inequality follows from the fact that u_0^* is the corresponding optimal solution of the leader's local optimal control problem from Lemma 5.2. Hence, from (5.33) and (5.34), we have

$$J_0^N(u_0^*, u^{N*}(u_0^*)) \leq \inf_{u_0 \in \mathcal{U}_0} J_0^N(u_0, u^{N*}(u_0)) + \epsilon_2,$$

where $\epsilon_2 = O(\frac{1}{\sqrt{N}} + \epsilon_N)$. This completes the proof. □

In Section 5.3.3 below Theorem 5.1, we had shown that for any $u_0 \in \mathcal{U}_0$ and when $N \to \infty$, the ϵ-Nash equilibrium in Theorem 5.1 is equivalent to the centralized Nash equilibrium obtained in Section 5.2.2 (that is the set of the centralized strategies u_i', $1 \leq i \leq N$, in (5.5)), provided that the solution of the FBSDE in (5.6) exists for all N. This implies that when $N \to \infty$, the (ϵ_1, ϵ_2)-Stackelberg equilibrium becomes identical to the centralized Stackelberg equilibrium that is a solution of the leader's optimization problem in (5.7), since the optimization problems for the leader in (5.7) and (5.22) become identical (note that $(1/N)\sum_{i=1}^{N} x_i'(t)$ in (5.7) converges to the mean field process $z(t)$ in (5.15) as $N \to \infty$). Therefore, in view of Theorems 5.1 and 5.2, the (ϵ_1, ϵ_2)-Stackelberg equilibrium is approximated centralized Stackelberg equilibrium with an approximation factor $O(\frac{1}{\sqrt{N}} + \epsilon_N)$. As a consequence, when there is a large number of followers, there is no need for the followers to share the information among themselves, which is of course less restrictive than the original Stackelberg game discussed in Section 5.2.2.

5.4.3 Optimality of Leader's Policy with State-feedback Policies by the Followers

We now consider the situation when the followers have, instead, access to state feedback information, in which case the leader is faced with an optimal control problem with additional constraints given in (5.25), instead of (5.24). For notational simplicity, we assume that $F = 0$, which implies that coupling effect between the leader and the followers appear in their cost functions only.

Similar to Lemma 5.2, we can then show that $(x_0^*, u_0^*) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n+p})$ is the corresponding optimal solution if and only if

$$u_0^*(t) = R_0^{-1} B_0^T p_0(t) - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^{K} \pi_k \bar{\lambda}_k(t) - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^{K} \pi_k Z_k(t) \bar{\xi}_k(t), \tag{5.35}$$

where Z_k is the RDE given in (5.12) and the set of corresponding FBSDEs is given by

$$
\begin{cases}
dx_0^*(t) = \Big[A_0 x_0^*(t) + B_0 R_0^{-1} B_0^T p_0(t) - B_0 R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k \bar{\lambda}_k(t) \\
\qquad\qquad - B_0 R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k Z_k(t) \bar{\xi}_k(t)\Big] dt + D_0 dW_0(t) \\
dp_0(t) = \Big[-A_0^T p_0(t) + Q_0(x_0^*(t) - H_0 z(t))\Big] dt + q_0 dW_0(t) \\
d\bar{\lambda}_k(t) = \Big[-\mathbb{G}_k^T(t)\bar{\lambda}_k(t) + H_0^T Q_0(H_0 z(t) - x_0(t)) + H^T Q \sum_{k=1}^K \pi_k \bar{\xi}_k(t)\Big] dt \\
\qquad\qquad + \bar{q}_k(t) dW_0(t) \\
d\bar{\xi}_k(t) = \Big[\mathbb{G}_k(t)\bar{\xi}_k(t) - BR^{-1}B^T \bar{\lambda}_k(t)\Big] dt \\
dz_k(t) = \Big[\mathbb{G}_k(t) z_k(t) + BR^{-1}B^T \bar{\phi}_k(t)\Big] dt - BR^{-1}L\Big[R_0^{-1} B_0^T p_0(t) \\
\qquad\qquad - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k \bar{\lambda}_k(t) \\
\qquad\qquad - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k Z_k(t) \bar{\xi}_k(t)\Big] dt \\
d\bar{\phi}_k(t) = \Big[-\mathbb{G}^T(t)\phi(t) - QHz(t)\Big] dt - Z_k(t) BR^{-1}L\Big[R_0^{-1} B_0^T p_0(t) \\
\qquad\qquad - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k \bar{\lambda}_k(t) \\
\qquad\qquad - R_0^{-1} L^T R^{-1} B^T \sum_{k=1}^K \pi_k Z_k(t) \bar{\xi}_k(t)\Big] dt \\
x_0^*(0) = x_0(0),\ \bar{\xi}_k(0) = 0,\ z_k(0) = \bar{x} \\
p_0(T) = 0,\ \bar{\lambda}_k(T) = 0,\ \bar{\phi}_k(T) = 0,\ k = 1, 2, ..., K.
\end{cases}
\tag{5.36}
$$

It is easy to show that $\lambda_k(t) = Z_k(t)\bar{\xi}_k(t) + \bar{\lambda}_k(t)$ for all $k \in \mathcal{K}$. By substituting this transformation into $\bar{\xi}_k$, we have $\bar{\xi}_k \equiv \xi_k$, which implies that u_0^* in (5.35) is equivalent to the optimal strategy of the leader when the followers are under the open-loop representation given in (5.31). Moreover, since the RDE Z_k has a unique solution as discussed in Section 5.3.1, the existence condition of the solution of the FBSDEs in (5.36) is the same as the one in Proposition 5.4. Therefore, when the leader announces (5.35) instead of (5.31) to the N followers, the set of the optimal decentralized strategies for the leader and the followers, that is, u_0^* in (5.35) and the set of u_i^*s in (5.14) with u_0^*, still constitute an (ϵ_1, ϵ_2)-Stackelberg equilibrium, where $\epsilon_1 = \epsilon_2 = O(\frac{1}{\sqrt{N}} + \epsilon_N)$.

We should mention that in standard linear-quadratic Stackelberg stochastic differential games, the information of the follower does not play an important role. That is, as discussed in [61], regardless of whether the follower has open-loop or closed-loop perfect state information, the Stackelberg strategy

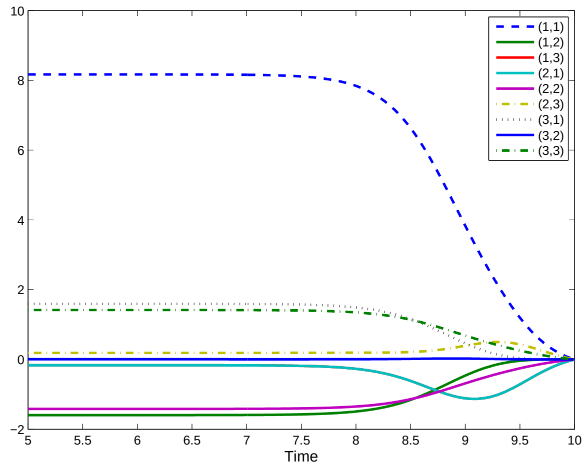


Figure 5.1: $\Lambda(t)$ over time with $\Lambda(10) = 0$.

of the leader is the same, since different information structures do not lead to different optimization problems for the leader. The results in this subsection also confirm the same phenomenon in mean field Stackelberg games.

5.5 Numerical Examples

This section provides numerical examples. We first consider the case of uniform followers ($K = 1$), and then discuss the heterogeneous case with $K = 4$.

5.5.1 The Case of Uniform Followers

Consider the SDE for $\mathcal{P}_0$ and $\mathcal{P}_i$, $1 \leq i \leq N$:

$$\begin{aligned} dx_0(t) &= [2x_0(t) + u_0(t)]dt + dW_0(t) \\ dx_i(t) &= [1.3x_i(t) + 2u_i(t)]dt + dW_i(t), \end{aligned}$$

and the performance indices

$$\begin{aligned} J_0 &= \mathbb{E}\int_0^{10} [(x_0(t) - 0.8x^N(t))^2 + 2u_0^2(t)]dt \\ J_i &= \mathbb{E}\int_0^{10} [(x_i(t) - 0.7x^N(t))^2 + 2u_0^2(t) + u_i(t)u_0(t)]dt, \end{aligned}$$

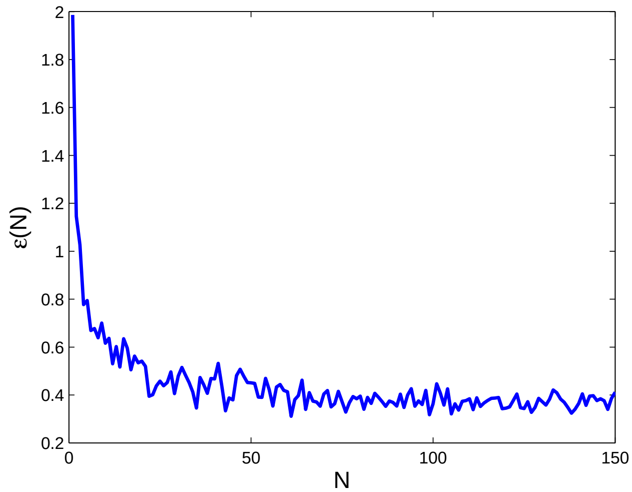


Figure 5.2: $\epsilon(N)$ with respect to N.

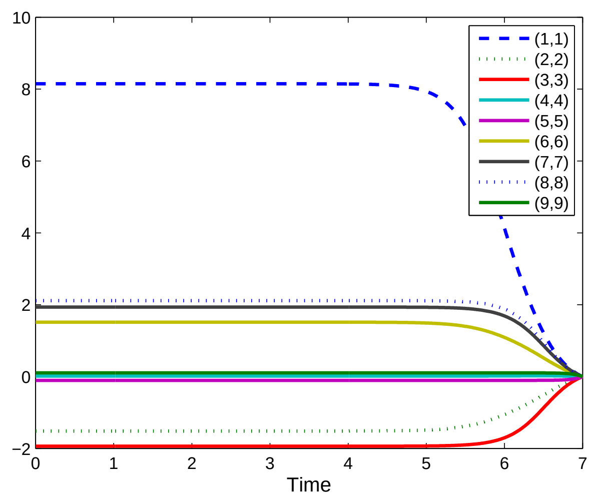


Figure 5.3: $\Lambda(t)$ over time with $\Lambda(7) = 0$ for the heterogeneous case.

where $x_0(0)$ and $x_i(0)$ are distributed according to $\mathcal{N}(0, 1)$. It can be checked that with above parameters, the LMI condition in Proposition 5.4 holds, where $\Lambda(t) \in \mathbb{R}^{3\times 3}$ is depicted in Fig 5.1. Note that $\Lambda(t)$ is neither symmetric nor positive semi-definite. Figure 5.2 plots $\epsilon(N) := (\mathbb{E}\int_0^{10} \|x^N(t) - z(t)\|^2 dt)^{1/2}$. Note that for the case of uniform followers, we have $\epsilon(N) = O(1/\sqrt{N})$. As can be seen, $\epsilon(N)$ converges as $N \to \infty$.

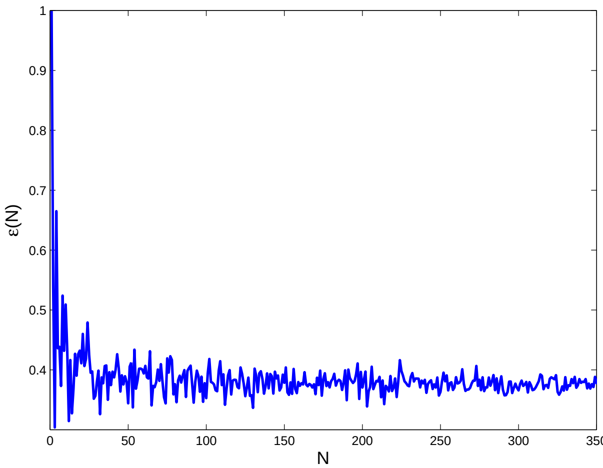


Figure 5.4: $\epsilon(N)$ with respect to N for the heterogeneous case.

5.5.2 The Heterogeneous Case with $K = 4$

We now consider the same model as in Section 5.5.1, except that $T = 7$ and there are $K = 4$ distinct models of the followers with

$$A(\theta_i) \in \{1.3,\ 1.7,\ 1.9,\ -3.8\},\ \pi = \begin{pmatrix} 2/4 & 1/8 & 1/8 & 1/4 \end{pmatrix}.$$

The nonsymmetric RDE Λ is a 9×9 matrix, whose diagonal terms are depicted in Figure 5.3. The plot of $\epsilon(N)$ is shown in Fig. 5.4. It also converges as $N \to \infty$.

5.6 Conclusions

In this chapter, we have considered mean field Stackelberg differential games, where there is one leader and a large number, say N, of heterogeneous followers with K distinct types. We have used the stochastic mean field approximation to characterize the best estimate of the actual mean field process of the followers. We have shown that with the approximated mean field process and each fixed strategy of the leader, the optimal decentralized controllers of the followers, which are solutions of the followers' local control problems, constitute an ϵ-Nash equilibrium, where ϵ converges to zero as $N \to \infty$. For the leader's problem, we have identified an LMI condition under which the leader's local optimal control problem admits a unique optimal decentralized controller. We have shown that the optimal decentralized controllers for the leader and the followers constitute an (ϵ_1, ϵ_2)-Stackelberg equilibrium, where ϵ_1 and ϵ_2 both converge to zero as $N \to \infty$.

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Recap

In this thesis, we studied various problems of control and estimation with limited information within a game-theoretic approach, to address communication constraints and structural constraints in modern control systems.

In the first part of the thesis, that is Chapters 2 and 3, we considered control and estimation problems over unreliable communication channels within the stochastic zero-sum dynamic game framework.

In Chapter 2, we obtained the stochastic minimax state estimator (SMSE) for the case when the estimator receives the sensor measurement intermittently governed by the Bernoulli process. We analyzed the asymptotic behavior of the estimation error in terms of a generalized stochastic Riccati equation (GSRE). In particular, we identified conditions on the rate of intermittent observations and the disturbance attenuation parameter, above which 1) the expected value of the GSRE is bounded from below and above by the deterministic quantities, and 2) the sequence of the norm generated by the GSRE converges weakly to a unique stationary distribution. Finally, we identified explicit lower and upper bounds for the expected value of the GSRE, and showed their convergence.

In Chapter 3, we considered the minimax control problem for TCP- and UDP-like unreliable communication channels that are modeled by two independent Bernoulli processes. We obtained the output feedback minimax controllers in both cases. We also characterized the set of threshold-type existence conditions for both problems in terms of the communication channel loss rates and the disturbance attenuation parameter, above which the minimax controller is able to achieve the desired control performance and stability. Finally, we identified a trade-off between disturbance attenuation and fundamental communication limitations.

In the second part of the thesis, that is Chapters 4 and 5, we considered two different large-scale optimization problems via mean field game theory to address structural constraints in the large-scale system.

In Chapter 4, risk-sensitive and robust mean field games were considered. We obtained an ϵ-Nash equilibrium for each corresponding problem, where the individual Nash strategies are decentralized as a function of local state information, and ϵ can be picked arbitrarily small when the number of agents is arbitrarily large. We showed that the two ϵ-Nash equilibria are *partially equivalent* in the sense that the individual Nash strategies for both problems share the same control law, but are determined by the different mean field system that provides the best estimation of the actual mean field behavior. Finally, we showed that both ϵ-Nash equilibria feature robustness in terms of the disturbance attenuation parameter.

The last chapter, Chapter 5, considered mean field Stackelberg differential games. We identified the approximated stochastic mean field behavior of the followers governed by the leader's strategy. With the approximated stochastic mean field behavior, we obtained a decentralized ϵ-Nash equilibrium for the followers. Moreover, the decentralized (ϵ_1, ϵ_2)-Stackelberg strategy of the leader was obtained by solving the leader's local nonstandard optimal control problem. We identified an existence condition for the (ϵ_1, ϵ_2)-Stackelberg strategy in terms of a linear matrix inequality. Finally, as a consequence of decentralization, we showed that ϵ_1 and ϵ_2 can be picked arbitrarily small when the number of followers is arbitrarily large.

6.2 Directions for Future Work

There are several interesting open problems on networked control systems and mean field games as natural outgrowths of the problems treated in this thesis. We identify some of these below.

Networked Control Systems

Chapter 2 studied the asymptotic behavior of the GSRE in the expectation sense and the weak convergence sense. The asymptotic behavior of the GSRE can also be analyzed in the probabilistic sense, as was done for the Kalman filtering problem in [50, 51]. In this case, one needs to show boundedness of $\mathbb{P}(\Sigma_k \leq M)$, where $M \geq 0$. This will provide a different perspective to the asymptotic behavior of the GSRE, which cannot be interpreted in the expectation sense.

In Chapters 2 and 3, the corresponding unreliable communication channel was taken to be temporally uncorrelated in terms of an i.i.d. Bernoulli process. This can be extended to the temporally correlated one with the Gilbert-Elliot channel model, in which case the communication channel can be regarded as a two-state irreducible and stationary Markov chain [27, 53]. The certainty equivalence principle for the TCP-case still holds. In order to obtain the corresponding minimax controller, we need to use dynamic programming. This, however, requires more involved steps than Section 3.3, since depending on the acknowledged information, two different zero-sum games appear at each time k to derive a minimax controller.

The minimax control problem for the TCP-case in Chapter 3 considers the case of perfect acknowledgments of control packet losses. In real-world systems, however, these acknowledgments may not be transmitted reliably due to the unreliable nature of the reverse channel. This unreliability can be modeled within a probabilistic framework, in which the controller receives acknowledgments perfectly with probability $\nu \in [0, 1]$. Note that when $\nu = 0$, the problem is equivalent to the UDP-case as studied in Chapter 3. This would be an interesting and challenging direction of research to pursue.

Finally, it would be interesting to study minimax control and estimation problems for nonlinear dynamical systems over unreliable communication channels, which is a problem that has been studied in [124] within the LQG framework.

Mean Field Games

The results obtained in Chapter 4 have shown how risk-sensitive optimal control, stochastic zero-sum differential game, and risk-neutral LQG control are related within the framework of mean field games. A further interesting problem would be to establish such relationships for systems described by general nonlinear stochastic differential equations under general cost functions. Moreover, it would also be interesting to study risk-sensitive and robust mean field games with major and minor players, as was done for the risk-neutral case in [84, 85, 86].

The analysis developed in Chapter 5 considers the one leader case. This can be extended to the problem of multiple leaders and followers. In this extension, the leaders will also play a Nash game among themselves; therefore, the problem can be viewed as a combination of Stackelberg and Nash games as mentioned in [118]. In this case, due to the multiple influence of the leaders on the mean field behavior of the followers, the leaders' local optimal control problems will be more involved than that in Chapter 5.

Another possible extension of the results of Chapter 5 would be to the case when the mean field coupling term is also included in the individual stochastic differential equations, which will introduce a more general coupling scenario between the followers and the leader.

One can also consider mean field Stackelberg dynamic games with discrete-time system dynamics for the leader and the followers. This problem can also be viewed as an extension of [125], where a similar problem is studied with state feedback information and a simplified cost function.

Finally, studying mean field games for singularly perturbed systems (SPSs) is an important area of research in control, since SPSs are able to capture general dynamic behavior by using two different time scales, "fast" and "slow," where the fast term corresponds to dynamics with small time constants. A major challenge in studying this problem is to construct a separation of time scales to design a local optimal controller that provides a best estimate mean field behavior, and to construct a corresponding mean field system.

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Proof of Convergence of the MGRE

This section provides proofs of convergence of (2.11) and the MGRE in (2.12). We rewrite the modified generalized Riccati equation (MGRE) in (2.12):

$$X_{k+1} = (1-\beta)h_1(\gamma, X_k) + \beta h_2(\gamma, X_k) =: h(\gamma, \beta, X_k), \tag{A.1}$$

where $X_0 = Q_0^{-1}$, and $h_1(\gamma, X)$ and $h_2(\gamma, \beta, X)$ are functions that are defined by

$$h_1(\gamma, X) := A(X^{-1} - \gamma^{-2}Q)^{-1}A^T + DD^T \tag{A.2a}$$

$$h_2(\gamma, X) := A(X^{-1} - \gamma^{-2}Q + C^T V^{-1} C)^{-1}A^T + DD^T. \tag{A.2b}$$

In terms of h_1 and h_2, the modified generalized algebraic Riccati equation (MGARE) can be rewritten as

$$X = (1-\beta)h_1(\gamma, X) + \beta h_2(\gamma, X) =: h(\gamma, \beta, X). \tag{A.3}$$

The following condition is needed:

$$\rho(XQ) < \gamma^2. \tag{A.4}$$

Before stating our main result on the convergence of (A.1), we observe that the MGARE in (A.3) is a convex combination of two different AREs. For the first part, consider the problem of maximizing

$$H_\gamma^1(p'_{0:N-1}) = |z'_N|^2_{Q_0^{-1}} + \sum_{k=0}^{N-1} |z'_k|^2_Q - \gamma^2 |p'_k|^2, \tag{A.5}$$

subject to $z'_{k+1} = Az'_k + Dp'_k$. The second part of (A.3) is related to the zero-sum dynamic game with the objective function

$$H_\gamma^2(q''_{0:N-1}, p''_{0:N-1}) = |z''_N|^2_{Q_0^{-1}} + \sum_{k=0}^{N-1} |z''_k|^2_Q + |q''_k|^2 - \gamma^2 |p''_k|^2, \tag{A.6}$$

and state dynamics $z''_{k+1} = Az''_k + C^T V^{-1/2} q''_k + Dp''_k$, where $q''_{0:N-1}$ is the minimizer and $p''_{0:N-1}$ is the maximizer. Then we can solve (A.5) and (A.6) by dynamic programming (under the strict concavity assumption, see [61]), the corresponding Riccati equations are, respectively, given by

$$X'_{k+1} = h_1(\gamma, X'_k),\ X'_0 = Q_0^{-1} \tag{A.7}$$

$$X''_{k+1} = h_2(\gamma, X''_k),\ X''_0 = Q_0^{-1}, \tag{A.8}$$

where we used the time-reverse notation in [126] and then replaced the triple (A, D, G) with the ordered triple (A^T, G^T, D^T) where $Q = G^T G$.

It was shown in [106] that if (A.7) and (A.8) have fixed points, i.e., $X' = h_1(\gamma, X')$ and $X'' = h_2(\gamma, X'')$, satisfying (A.4) where $X', X'' \in \mathbb{S}^n_{>0}$, then they converge to X' and X'', respectively. This follows from monotonicity and continuity of (A.7) and (A.8) with the controllability and observability assumptions [126, 106].

We now show convergence of (A.1).

Lemma A.1. *Suppose that $X_0 \leq DD^T$, $\beta \in [0,1]$, and $\gamma > 0$ is fixed. Suppose that MGRE satisfies (A.4) for all k. If $\bar{X} \in \mathbb{S}^n_{>0}$ is a fixed point of (A.3) that satisfies (A.4), then as $k \to \infty$, $\{X_k\} \to X^+$ where X^+ is a fixed point of (A.3) satisfying (A.4).*

Proof. First note that when $X_0 = 0$, $\{X_k\}$ is a monotone sequence when $\rho(X_k Q) < \gamma^2$. To see this, at $k = 0$, we have $X_0 \leq h(\gamma, \beta, X_0) = X_1$. Now, assume that we have $X_{k-1} \leq X_k$. Then

$$\begin{aligned} X_k &= h(\gamma, \beta, X_{k-1}) \\ &= (1-\beta) h_1(\gamma, X_{k-1}) + \beta h_2(\gamma, X_{k-1}) \\ &\leq (1-\beta) h_1(\gamma, X_k) + \beta h_2(\gamma, X_k) \\ &= h(\gamma, \beta, X_k) = X_{k+1}, \end{aligned}$$

where the inequality follows from Lemma A.3(viii). Therefore, if there exists

$M > 0$ such that $M \geq X_k$ for all k satisfying (A.4), then $\{X_k\}$ converges to X^+, and by continuity, it is a fixed point of the MGARE satisfying (A.4).

We next show that if there is a fixed point $\bar{X} \in \mathbb{S}^n_{>0}$ of the MGARE satisfying (A.4), then $\bar{X} \geq X_k$ for all k. Clearly, at $k = 0$, we must have $\bar{X} \geq X_0$. Now, suppose we have $\bar{X} - X_k \geq 0$. Consider

$$
\begin{aligned}
&\bar{X} - X_{k+1} \\
&= (1-\beta)\big(A((\bar{X}^{-1} - \gamma^{-2}Q)^{-1} - (X_k^{-1} - \gamma^{-2}Q)^{-1})A^T\big) \\
&\quad + \beta\big(A((\bar{X}^{-1} - \gamma^{-2}Q + C^T V^{-1} C)^{-1} - (X_k^{-1} - \gamma^{-2}Q + C^T V^{-1} C)^{-1})A^T\big).
\end{aligned}
$$

Then it is easy to see that $\bar{X} \geq X_{k+1}$. Therefore, $\bar{X} \geq X_k$ for all k.

Finally, if $X_0 = 0$ and there is a fixed point $\bar{X}$ satisfying (A.4), then due to the monotonicity of the MGRE, $\{X_k\}$ converges to X^+ as $k \to \infty$, where by continuity X^+ is a fixed point of (A.3) satisfying (A.4). The existence of the fixed point of the MGARE is shown in the next lemma. Note that since the monotonicity condition holds with $X_0 \leq DD^T$, we have convergence for all initial conditions $X_0 \leq DD^T$. This completes the proof of the lemma. $\square$

We now show existence of the fixed point of the MGARE (A.3) that satisfies (A.4).

Lemma A.2. *There is a finite $\gamma > 0$ and a fixed point of the MGARE in $\mathbb{S}^n_{\geq 0}$ that satisfies (A.4) for all $\beta \in (0, 1]$.*

Proof. First assume that A is stable, and observe that (A.2a) and (A.2b) are standard AREs; hence under the controllability and observability assumption, each of them has a unique fixed point in $\mathbb{S}^n_{\geq 0}$ that satisfies (A.4) for some γ_1 and γ_2, respectively, due to [106]. Choose $\gamma := \max\{\gamma_1, \gamma_2\}$ and let X_1 and X_2 be fixed points of (A.2a) and (A.2b), respectively. Clearly, we have $X_2 \leq X_1$. Now, consider the function of the convex combination of (A.2a) and (A.2b), namely (A.3). When $\beta = 0$, its fixed point is X_1, and X_2 when $\beta = 1$. Let $\mathcal{E} = \{X : X_2 \leq X \leq X_1\}$. Then (A.3) can be restricted to a continuous mapping of $\mathcal{E}$ into itself due to Lemma A.3(i). Moreover, $\mathcal{E}$ is convex, compact and nonempty subset of $\mathbb{S}^n_{\geq 0}$. Then existence follows from Brouwer's fixed point theorem [61].

Now, we consider the case when A is unstable, in which case (A.2a) does not have a fixed point for all γ, but (A.2b) does for some γ, which satisfies

(A.4) [106]. Define

$$\beta'(\gamma) := \inf\{\beta \in (0,1) \ : \ \text{(A.3) has a fixed point that satisfies (A.4)}\}.$$

Then from Lemma A.3(i), for any $\beta > \beta'(\gamma)$, (A.3) can be restricted to a continuous mapping of $\mathcal{E}'$ into itself where $\mathcal{E}' = \{X : X_2 \leq X \leq X_3\}$ with $X_3 = h(\gamma, \beta, X_3)$. Then existence follows again from Brouwer's fixed point theorem. This completes the proof. □

Proposition 2.2. Note that for any $\beta > (1 - 1/\rho^2(A))$, we have $\sqrt{1-\beta}\rho(A) < 1$. Let $\tilde{A} := \sqrt{1-\beta}A$ and $\tilde{D} := \sqrt{1-\beta}D$. By Lemma A.2, there exist γ and $\breve{\Sigma}^+$ satisfying $\breve{\Sigma}^+ = h_1(\gamma, \breve{\Sigma}^+)$, and the condition $\rho(\breve{\Sigma}^+ Q) < \gamma^2$ with $\tilde{A}$ and $\tilde{D}$. Then convergence of $\{\breve{\Sigma}_k\}$ is equivalent to convergence of (A.7). This completes the proof. □

Proposition 2.3. Suppose A is stable. Then, for all $\beta \in [0, 1]$, there exists a solution to (2.14) that satisfies $\rho(\bar{\Sigma}Q) < \gamma^2$ due to Lemma A.2. Consequently, we have convergence due to Lemma A.1. Now, we consider the case when A is unstable. First observe that $\bar{\Gamma}(0)$ is empty. From the definitions, there exists a matrix $\bar{\Sigma}$ that solves (2.14) and satisfies $\rho(\bar{\Sigma}Q) < \gamma^2$. This existence is guaranteed due to Lemma A.2. Then, convergence follows from Lemma A.1. This completes the proof. □

Proposition 2.4. (i) The result follows from Lemma A.3(i) in Appendix A.2 since for a given γ, we have $\rho(\bar{\Sigma}(\gamma, \beta_1)Q) \leq \rho(\bar{\Sigma}(\gamma, \beta_2)Q) \leq \gamma^2$, which leads to $\bar{\Gamma}(\beta_2) \subseteq \bar{\Gamma}(\beta_1)$. The second statement can be shown in a similar manner.

(ii) The first statement follows from Lemma A.3(ii) because $\rho(\bar{\Sigma}(\gamma_1, \beta)Q) \leq \rho(\bar{\Sigma}(\gamma_2, \beta)Q) \leq \gamma_2^2 \leq \gamma_1^2$ shows that $\bar{\Lambda}(\gamma_2) \subseteq \bar{\Lambda}(\gamma_1)$. For the second statement, by Lemma A.3(vi), (2.14) converges to (A.10) as $\gamma \to \infty$.

□

A.2 Properties of the MGARE

The following lemma provides some useful properties of the MGARE in (A.3).

Lemma A.3. *Suppose that for a fixed $\gamma > 0$, $X \in \mathbb{S}^n_{>0}$ satisfies (A.4), and $Q = G^T G \in \mathbb{S}^n_{\geq 0}$. Then the following are true:*

(i) If $\beta_2 \geq \beta_1$, then $h(\gamma, \beta_1, X) \geq h(\gamma, \beta_2, X)$.

(ii) If $\gamma_2 \geq \gamma_1$, then $h(\gamma_1, \beta, X) \geq h(\gamma_2, \beta, X)$.

(iii) Suppose $\rho \in [0, 1]$ and $X := \rho X_1 + (1-\rho) X_2$ where $X_1, X_2 \in \mathbb{S}^n_{>0}$. Then $h_1(\gamma, X) \leq \rho h_1(\gamma, X_1) + (1-\rho) h_1(\gamma, X_2)$.

(iv) $h_2(\gamma, X) = \min_U \max_L \varsigma'(X, U, L) = \max_L \min_U \varsigma'(X, U, L)$, where

$$\begin{aligned}\varsigma'(X, U, L) = {} & (A + LG + UV^{-1/2}C) X (A + LG + UV^{-1/2}C)^T \\ & - \gamma^2 LL^T + UU^T.\end{aligned} \tag{A.9}$$

(v) We have the following inequalities:

$$(1-\beta) AXA^T + DD^T \leq (1-\beta) h_1(\gamma, X) \leq h(\gamma, \beta, X).$$

(vi) As $\gamma \to \infty$, (A.3) can be written as

$$h(\gamma, \beta, X) = AXA^T + DD^T - \beta AXC^T (CXC^T + V)^{-1} CXA^T. \tag{A.10}$$

(vii) As $\gamma \to \infty$, $h_1(\gamma, X) = (1-\beta)(AXA^T + DD^T)$.

(viii) For any $X_1 \leq X_2$ satisfying (A.4), we have $h_1(\gamma, X_1) \leq h_1(\gamma, X_2)$, $h_2(\gamma, X_1) \leq h_2(\gamma, X_2)$, and $h(\gamma, X_1) \leq h(\gamma, X_2)$.

Proof. (i) Consider $(\beta_2 - \beta_1) h_1(\gamma, X) - (\beta_2 - \beta_1) h_2(\gamma, X)$. Since $V \in \mathbb{S}^n_{>0}$ and C cannot be a zero matrix, under (A.4), $h_1(\gamma, X) > h_2(\gamma, X)$; thus, completing the proof.

(ii) The result follows by inspection.

(iii) Define

$$\varsigma(X, K) := (A + KG) X (A + KG)^T + DD^T - \gamma^2 KK^T. \tag{A.11}$$

By applying the matrix inversion lemma, we have

$$
\begin{aligned}
h_1(\gamma, X) &= AXA^T + AXG^T(\gamma^2 I - GXG^T)^{-1}GXA^T + DD^T \\
&= \max_K \varsigma(X, K),
\end{aligned}
$$

where the last equality is achieved by $K_X = AXG^T(\gamma^2 I - GXG^T)^{-1}$, since the above optimization problem is quadratic and concave in K. Now, we have

$$
\begin{aligned}
h_1(\gamma, X) &= \varsigma(X, K_X) \\
&= \rho\varsigma(X_1, K_X) + (1-\rho)\varsigma(X_2, K_X) \\
&\le \rho\varsigma(X_1, K_{X_1}) + (1-\rho)\varsigma(X_2, K_{X_2}) \\
&= \rho h_1(\gamma, X_1) + (1-\rho) h_1(\gamma, X_2).
\end{aligned}
$$

(iv) Let us consider the dual problem, i.e., the quadruple $(A, V^{-1/2}C, D, G)$ is replaced with the ordered quadruple (A^T, B^T, G^T, D^T). Then by using the matrix inversion lemma,

$$
\begin{aligned}
h_2'(\gamma, X) &= A^T(X^{-1} - \gamma^{-2}DD^T + BB^T)^{-1}A + Q \\
&= F_{h_2'}^T X F_{h_2'} + Q + P_{h_2'}^{1,T} P_{h_2'}^1 - \gamma^2 P_{h_2'}^{2,T} P_{h_2'}^2,
\end{aligned}
$$

where

$$
\begin{aligned}
F_{h_2'} &= A + DP_{h_2'}^2 - BP_{h_2'}^1 \\
P_{h_2'}^1 &= (I + B^T(I + XD(\gamma^2 I - D^T XD)^{-1}D^T)XB)^{-1} \\
&\quad \times B^T(I + XD(\gamma^2 I - D^T XD)^{-1}D^T)XA \\
P_{h_2'}^2 &= (\gamma^2 I - D^T(I - XB(I + B^T XB)^{-1}B^T)XD)^{-1} \\
&\quad \times D^T(I - XB(I + B^T XB)^{-1}B^T)XA.
\end{aligned}
$$

It can be shown that

$$
h_2'(\gamma, X) = \min_{U'} \max_{L'} \varsigma''(X, U', L') = \max_{L'} \min_{U'} \varsigma''(X, U', L'),
$$

where

$$\varsigma''(X, U', L')$$
$$= (A + DL' + BU')^T X (A + DL' + BU') - \gamma^2 L'^{,T} L' + U'^{,T} U',$$

where the corresponding minimizer and maximizer can be written as

$$U'_X = -P^1_{h'_2},\ L'_X = P^2_{h'_2}.$$

Then the result follows by replacing the quadruple (A, B, G, D) with the ordered quadruple $(A^T, C^T V^{-1/2}, D^T, G^T)$, in which case the minimizer is $U_X = U'^{,T}_X$ and the maximizer is $L_X = L'^{,T}_X$ with the ordered quadruple $(A^T, C^T V^{-1/2}, D^T, G^T)$.

(v) Inequalities follow by inspection.

(vi) It can be shown by using the matrix inversion lemma.

(vii) It can be shown by using the matrix inversion lemma.

(viii) Notice that (A.11) and (A.9) are affine in X. Consider

$$h_1(\gamma, X_1) = \varsigma(X_1, K_{X_1}) \leq \varsigma(X_2, K_{X_1}) \leq \varsigma(X_2, K_{X_2}) = h_1(\gamma, X_2),$$

and

$$\begin{aligned} h_2(\gamma, X_1) &= \varsigma'(X_1, U_{X_1}, L_{X_1}) \\ &\leq \varsigma'(X_1, U_{X_2}, L_{X_1}) \\ &\leq \varsigma'(X_2, U_{X_2}, L_{X_1}) \leq \varsigma'(X_2, U_{X_2}, L_{X_2}) = h_2(\gamma, X_2). \end{aligned}$$

Hence, the result follows. This completes the proof of the lemma.

□

We now show that there is a class of γ and β such that $h(\gamma, \beta, X)$ is concave in X. Note that when $\beta = 0$, $h(\gamma, \beta, X)$ cannot be concave due to Lemma A.3(iii). We have the following relation:

$$\begin{aligned} h(\gamma, \beta, X) &= (1 - \beta) h_1(\gamma, X) + \beta h_2(\gamma, X) \\ &= (1 - \beta)\varsigma(X, K_X) + \beta \varsigma'(X, U_X, L_X), \end{aligned}$$

where $\varsigma(X, K)$ and $\varsigma'(X, U, L)$ are defined in (A.11) and (A.9), respectively, and K_X, L_X, and U_X are given in the proof of Lemma A.3(iii) and (iv).

Define

$$\Upsilon(X, K, U, L) := (1-\beta)\varsigma(X, K) + \beta\varsigma'(X, U, L).$$

Note that by definition, $\Upsilon(X, K, U, L)$ is affine in X. We then have

$$\begin{aligned} h(\gamma, \beta, X) &= \Upsilon(X, K_X, U_X, L_X) \\ &= \min_U \max_{K,L} \Upsilon(X, K, U, L) = \max_{K,L} \min_U \Upsilon(X, K, U, L), \end{aligned}$$

which follows because the zero-sum game is strictly convex in U and strictly concave in L and K, and $\varsigma(X, K)$ (resp. $\varsigma'(X, U, L)$) is independent of U and L (resp. K).

Let $X = \rho X_1 + (1-\rho)X_2$ where $\rho \in [0, 1]$. Consider

$$\begin{aligned} h(\gamma, \beta, X) &= \Upsilon(X, K_X, U_X) \\ &= \rho\Upsilon(X_1, K_X, U_X, L_X) + (1-\rho)\Upsilon(X_2, K_X, U_X, L_X) \\ &\geq \rho\Upsilon(X_1, K, U_X, L) + (1-\rho)\Upsilon(X_2, K, U_X, L), \ \forall K, L, \end{aligned}$$

where the inequality follows from the definition of the saddle point [61]. Let $K = K_{X_1}$ and $L = L_{X_1}$. Clearly, we have

$$\begin{aligned} h(\gamma, \beta, X) &\geq \rho\Upsilon(X_1, K_{X_1}, U_X, L_{X_1}) + (1-\rho)\Upsilon(X_2, K_{X_1}, U_X, L_{X_1}) \\ &\geq \rho\Upsilon(X_1, K_{X_1}, U_{X_1}, L_{X_1}) + (1-\rho)\Upsilon(X_2, K_{X_1}, U_{X_2}, L_{X_1}), \end{aligned}$$

since U_{X_1} and U_{X_2} are the corresponding minimizers. On the other hand, we also have the following relation due to the definition of the saddle point:

$$\begin{aligned} h(\gamma, \beta, X) &\leq \rho\Upsilon(X_1, K_X, U_{X_1}, L_X) + (1-\rho)\Upsilon(X_2, K_X, U_{X_1}, L_X) \\ &\leq \rho\Upsilon(X_1, K_{X_1}, U_{X_1}, L_{X_1}) + (1-\rho)\Upsilon(X_2, K_{X_2}, U_{X_1}, L_{X_2}). \end{aligned}$$

Note that by Lemma A.3(iii) and (iv),

$$\Upsilon(X_2, K_{X_1}, U_{X_2}, L_{X_1}) \leq \Upsilon(X_2, K_{X_2}, U_{X_2}, L_{X_2}),$$

and

$$\Upsilon(X_2, K_{X_2}, U_{X_1}, L_{X_2}) \geq \Upsilon(X_2, K_{X_2}, U_{X_2}, L_{X_2}).$$

Hence, it can be easily seen that we can choose γ and β such that for all $\rho \in [0, 1]$,

$$\begin{aligned} h(\gamma, \beta, X) &\geq \rho\Upsilon(X_1, K_{X_1}, U_{X_1}, L_{X_1}) + (1-\rho)\Upsilon(X_2, K_{X_2}, U_{X_2}, L_{X_2}) \quad \text{(A.12)} \\ &= \rho h(\gamma, \beta, X_1) + (1-\rho)h(\gamma, \beta, X_2). \end{aligned}$$

Then we have the desired result. The following lemma states the above discussion.

Lemma A.4. *Suppose that $X = \rho X_1 + (1-\rho)X_2 > 0$ where $\rho \in [0, 1]$ and X satisfies (A.4). Then $h(\gamma, \beta, X)$ is concave in X if one of the following conditions holds:*

(i) γ is sufficiently large.

(ii) $C^T V^{-1} C - \gamma^{-2} Q \geq 0$ and $\beta = 1$.

(iii) Equation (A.12) holds for some γ and β. □

Note that part (i) is the case considered in [22]. Part (ii) can be shown in a similar way to that in Lemma A.3(iii).

A.3 Kalman Filtering with Intermittent Observations

In this appendix, we recall the results on Kalman filtering with intermittent observations in [22, 46, 45].

Suppose that x_0 is a Gaussian random vector with zero mean and covariance matrix Q_0^{-1}, and $\{w_k\}$ and $\{v_k\}$ are i.i.d. Gaussian processes with zero mean and covariance matrices DD^T and V, respectively. Moreover, suppose that $(x_0, \{w_k\}, \{v_k\})$ are independent of each other. Suppose $\mathbb{P}(\beta_k = 1) = \beta = \lambda$. Then under the full information structure in (2.2), the Kalman filter and the associated stochastic Riccati equation (SRE) can be written as [22]

$$\hat{x}_{k+1} = A\hat{x}_k + \beta_k A P_k C^T (C P_k C^T + V)^{-1} (y_k - C\hat{x}_k) \tag{A.13}$$

$$P_{k+1} = A P_k A^T + DD^T - \beta_k A P_k C^T (C P_k C^T + V)^{-1} C P_k A^T \tag{A.14}$$

$$= A(P_k^{-1} + \beta_k C^T V^{-1} C)^{-1} A^T + DD^T,\ P_0 = Q_0^{-1}, \tag{A.15}$$

where $P_k := \mathbb{E}\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T \ |\mathcal{I}_k\}$ is an error covariance matrix and (A.15) is obtained by applying the matrix inversion lemma to (A.14). The error covariance matrix (A.14) (or (A.15)) was studied extensively in the literature. We state below the relevant result obtained by [22, 44, 46, 45].

Result A.1 (Boundedness of $\mathbb{E}\{P_k\}$)**.** *(i) There exists a critical value $\lambda_c \in [0, 1)$ such that*

$$\forall P_0, k,\ \mathbb{E}\{P_k\} \le M_{P_o} \text{ if } \lambda > \lambda_c$$
$$\exists P_0,\ \lim_{k\to\infty} \mathbb{E}\{P_k\} = \infty \text{ if } \lambda \le \lambda_c,$$

where $M_{P_o} \in \mathbb{S}^n_{\ge 0}$ depends on the initial condition of (A.14).

(ii) Define

$$S_{k+1} = (1-\lambda) A S_k A^T + DD^T =: v_1(S_k)$$
$$F_{k+1} = A F_k A^T + DD^T - \lambda A F_k C^T (C F_k C^T + V)^{-1} C F_k A^T =: v_2(F_k),$$

where $S_0 = F_0 = P_0$. Let $\breve{\lambda}$ and $\bar{\lambda}$ be such that

$$\lim_{k\to\infty} S_k = S \in \mathbb{S}^n_{\ge 0},\ \forall \lambda > \breve{\lambda} := 1 - \frac{1}{\rho(A)^2}$$
$$\lim_{k\to\infty} F_k = F \in \mathbb{S}^n_{\ge 0},\ \forall \lambda > \bar{\lambda},$$

where $S = v_1(S)$ and $F = v_2(F)$. Then λ_c satisfies $\breve{\lambda} \le \lambda_c \le \bar{\lambda}$. Moreover, if $\lambda > \bar{\lambda}$, then $S \le \mathbb{E}\{P_k\} \le F$ for all k.

(iii) The critical value satisfies $\lambda_c = \breve{\lambda}$ if one of the following three conditions holds:

(a) C is invertible or invertible on the observable subspace

(b) A has only one unstable eigenvalue

(c) The eigenvalues of A have distinct absolute values. □

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Properties of the GARE

Here, we provide some useful properties of the GARE in (3.13). For notational convenience, the "overbar" is dropped.

Lemma B.1. *Suppose* (A, B) *is controllable and* $(A, Q^{1/2})$ *is observable. Assume that given* γ *and* α*, (3.18) holds for all* k*. Assume further that the GARE has a solution* $Z := Z(\gamma, \alpha) \geq 0$ *which satisfies (3.16). Then:*

(i) $\tilde{Z}_k \leq \tilde{Z}_{k+1}$ *for all* k*.*

(ii) $Z \geq \tilde{Z}_k$ *for all* k*.*

Proof. (i) Note that $\tilde{Z}_k$ with $\alpha = 1$ is the GRE of the deterministic H^∞ optimal control problem, and its monotonicity was proven in [106]. Moreover when $\alpha = 0$, the GRE can be obtained by solving the following optimization problem:

$$\max_{w_{0:N-1}} |x_N|^2_{Q_N} + \sum_{k=0}^{N-1} |x_k|^2_Q - \gamma^2 |w_k|^2, \tag{B.1}$$

with the constraint of $x_{k+1} = Ax_k + Dw_k$. Then under the concavity condition, the monotonicity holds [106]. To prove the general case, note that

$Z_N = 0 \leq Z_{N-1}$. By induction, suppose $Z_{k+1} \geq Z_{k+2}$. Then we have

$$\begin{aligned}
V_k(x) &= |x|^2_{Z_k} \\
&= \min_u \max_w \Big[(1-\alpha)\mathbb{E}\Big\{V_{k+1}(Ax+Dw) + |x|^2_Q - \gamma^2|w|^2\Big|\mathcal{I}_k\Big\} \\
&\quad + \alpha\mathbb{E}\Big\{V_{k+1}(Ax+Bu+Dw) + |x|^2_Q + |u|^2_R - \gamma^2|w|^2\Big|\mathcal{I}_k\Big\}\Big] \\
&\geq \min_u \max_w \Big[(1-\alpha)\mathbb{E}\Big\{V_{k+2}(Ax+Dw) + |x|^2_Q - \gamma^2|w|^2\Big|\mathcal{I}_{k+1}\Big\} \\
&\quad + \alpha\mathbb{E}\Big\{V_{k+2}(Ax+Bu+Dw) + |x|^2_Q + |u|^2_R - \gamma^2|w|^2\Big|\mathcal{I}_{k+1}\Big\}\Big] \\
&= |x|^2_{Z_{k+1}} = V_{k+1}(x).
\end{aligned}$$

Here $V_k(x) = \mathbb{E}\{x^T Z_k x|\mathcal{I}_k\}$ is the saddle-point value of the dynamic game with only $N-k$ stages (see Lemma 3.1). Hence, we have $Z_k \geq Z_{k+1}$ for all k. Then the result follows by reversing the time index.

(ii) Note that when $\alpha = 1$ or $\alpha = 0$, $Z \geq \tilde{Z}_k$ for all k, which was shown in [106]. To see the general case, when $k = 0$, $Z \geq \tilde{Z}_0 = 0$. Suppose $Z - \tilde{Z}_{k+1} \geq 0$. Under this assumption and the fact that α is positive, it can be checked that $Z - \tilde{Z}_k \geq 0$. Therefore, the result follows. □

B.2 Certainty Equivalence Principle

The main idea of the certainty equivalence principle for H^∞ control is as follows ([106]): *At time k, given the state feedback minimax controller, one should first look for the worst past disturbances that maximize the cost function under the specified information structure and then find the worst-case state estimator, $\hat{x}_k$, that corresponds to the worst past disturbances. If such an $\hat{x}_k$ exists, then the minimax controller can use it in place of the state x_k to generate the control action.*

In this appendix, we show that the TCP problem formulated in Section 3.2 in Chapter 3 satisfies the three basic properties of the certainty equivalence principle in [106], but the UDP problem does not.

Toward this end, we use the notation $s := \{s_k\} \in \mathbf{S}'$ and $s^\tau := \{s_k\}_{k=0}^{\tau} \in \mathbf{S}^\tau$. Let the set of disturbances Ω be $(x_0, w) =: \omega \in \Omega := \mathbb{R}^n \times \mathbf{W}'$. For the LTI system, let the solutions of (3.1a) and (3.1b) be $x_t = \phi_t(u, w, x_0, \{\alpha_k\}_{k=0}^{t-1})$ and $y_t = \eta_t(u, w, x_0, \{\alpha_k\}_{k=0}^{t-1}, \beta_t)$. By using the inherent causality, we have

$x_t = \phi_t(u^{t-1}, w^{t-1}, x_0, \{\alpha_k\}_{k=0}^{t-1})$ and $y_t = \eta_t(u^{t-1}, w^t, x_0, \{\alpha_k\}_{k=0}^{t-1}, \beta_t)$. Let Θ^τ be the set of realizations of packet drops until $\tau \in [0, N-1]$. Now, for any given $\tau \in [0, N-1]$ and $(\bar{u}, \bar{y}, \bar{\kappa}) \in \mathbf{U}^\tau \times \mathbf{Y}^\tau \times \Theta^\tau$, we define the following subset Ω_τ of Ω:

$$\Omega_\tau(\bar{u}, \bar{y}, \bar{\kappa}) := \{\omega \in \Omega \ : \ \eta_k(\bar{u}, \omega, \bar{\kappa}) = \bar{y}_k, \ k = 0, 1, ..., \tau\},$$

where the set is compatible with all disturbance sequences in Ω. We also introduce the following set which is the set of restrictions of the elements of Ω_τ to $[0, \tau]$:

$$\Omega_\tau^\tau(\bar{u}, \bar{y}, \bar{\kappa}) := \{\omega^\tau \in \Omega^\tau : \omega \in \Omega_\tau(\bar{u}, \bar{y}, \bar{\kappa})\}.$$

Note that Ω_τ and Ω_τ^τ are the sets that are related to the disturbances, which are compatible with observed sequences of control, measurement, and realizations of packet drops.

Now, it can be shown that the information process $(u, \omega, \kappa) \mapsto \{\Omega_\tau\}$ carries *consistent*, *perfect recall*, and *nonanticipative* properties introduced in [106, page 249]. Hence, we are now in a position to apply the certainty equivalence principle for the TCP problem. In particular, the zero-sum dynamic game for the TCP-case formulated in Section 3.2 in Chapter 3 can be studied through the following three steps:

(a) State feedback minimax control by assuming that the controller has the actual state information.

(b) Minimax estimation under the TCP-like information structure.

(c) Synthesis of the results in (a) and (b) by characterizing the worst-case state estimator, say $\hat{x}_k$, that will be used in the minimax controller obtained in part (a) by replacing the true state with $\hat{x}_k$.

For the UDP-case, on the other hand, due to the absence of acknowledgments, we cannot construct the above information process. Therefore, the zero-sum dynamic game of the UDP-case cannot be solved by the certainty equivalence principle.

APPENDIX C

APPENDIX FOR CHAPTER 4

C.1 Preliminary Results

Lemma C.1. *(i) The following inequality holds:*

$$\begin{aligned}
&\int_0^T \|x_i^* - f_N^*\|_Q^2 + \|u_i^*\|_R^2 dt \\
&\leq \int_0^T \|x_i^* - g^*\|_Q^2 + \|u_i^*\|_R^2 dt + \|Q\| \int_0^T \|f_N^* - g^*\|^2 dt \\
&\quad + 2\|Q\| \left(\int_0^T \|x_i^* - g^*\|^2 dt \right)^{\frac{1}{2}} \left(\int_0^T \|g^* - f_N^*\|^2 dt \right)^{\frac{1}{2}}.
\end{aligned}$$

(ii) The following holds:

$$\begin{aligned}
&\int_0^T \|x_i - g^*\|_Q^2 + \|u_i\|_R^2 dt \\
&\leq \int_0^T \|x_i - f_N^{-i^*}\|_Q^2 + \|u_i\|_R^2 dt + |F_1| + |F_2|,
\end{aligned}$$

where

$$\begin{aligned}
F_1 &\triangleq 2\|Q\| \left(\int_0^T \|x_i - g^*\|^2 dt \right)^{\frac{1}{2}} \left(\int_0^T \|g^* - f_N^*\|^2 dt \right)^{\frac{1}{2}} \\
F_2 &\triangleq \frac{2\|Q\|}{N} \left(\int_0^T \|x_i - g^*\|^2 dt \right)^{\frac{1}{2}} \left(\int_0^T \|x_i^* - x_i\|^2 dt \right)^{\frac{1}{2}}.
\end{aligned}$$

Proof. (i) It can be shown by using the following relation:

$$\begin{aligned}
&\int_0^T \|x_i^* - f_N^*\|_Q^2 + \|u_i^*\|_R^2 dt \\
&\leq \int_0^T \|x_i^* - g^*\|_Q^2 + \|u_i^*\|_R^2 dt + \int_0^T \|f_N^* - g^*\|_Q^2 dt \\
&\quad + 2\left(\int_0^T \|x_i^* - g^*\|_Q^2 dt\right)^{\frac{1}{2}} \left(\int_0^T \|g^* - f_N^*\|_Q^2 dt\right)^{\frac{1}{2}}.
\end{aligned}$$

(ii) We have (a.s.)

$$\int_0^T \|x_i - f_N^{-i^*}\|_Q^2 + \|u_i\|_R^2 dt \geq \int_0^T \|x_i - g^*\|_Q^2 + \|u_i\|_R^2 dt + U_1 + U_2,$$

where

$$\begin{aligned}
U_1 &= 2\int_0^T (x_i - g^*)^T Q (g^* - f_N^*) dt \\
U_2 &= \frac{2}{N}\int_0^T (x_i - g^*)^T Q (x_i^* - x_i) dt.
\end{aligned}$$

Now,

$$\begin{aligned}
|U_1| &\leq 2\|Q\| \left(\int_0^T \|x_i - g^*\|^2 dt\right)^{\frac{1}{2}} \left(\int_0^T \|g^* - f_N^*\|^2 dt\right)^{\frac{1}{2}} \\
|U_2| &\leq \frac{2\|Q\|}{N} \left(\int_0^T \|x_i - g^*\|^2 dt\right)^{\frac{1}{2}} \left(\int_0^T \|x_i^* - x_i\|^2 dt\right)^{\frac{1}{2}},
\end{aligned}$$

and this proves the lemma. □

Lemma C.2. *(i) Suppose that Assumptions 4.1 and 4.2 hold. Then we have $\mathcal{T}(x) \in \mathcal{C}_n^b$ for any $x \in \mathcal{C}_n^b$.*

(ii) $\bar{x}_\theta(t)$ in (4.19) is equicontinuous and uniformly bounded on $\Theta \times X$.

Proof. (i) The result follows from [127, Lemma 9.1].

(ii) The boundedness follows from the compactness of Θ and X. As for equicontinuity, see [71, Lemma 5.1].

□

Lemma C.3. *Under the conditions in Proposition 4.4, the closed-loop system (4.22) with the optimal controller in (4.25) and the worst-case disturbance in (4.26) satisfies* $\mathbb{E}\{\|x(T)\|^2\} = o(T)$ *and* $\mathbb{E}\{\int_0^T \|x(t)\|^2 dt\} = O(T)$.

Proof. We first prove $\mathbb{E}\{\int_0^T \|x(t)\|^2 dt\} = O(T)$. Consider

$$\begin{aligned}\mathbb{E}\{\|x(t)\|^2\} &\leq 8\|e^{Ht}\bar{x}\|^2 + 8\Big\|\int_0^t e^{H(t-\tau)}BR^{-1}B^T s(\tau)d\tau\Big\|^2 \\ &\quad + 8\mathbb{E}\Big\{\Big\|\sqrt{\mu}\int_0^t e^{H(t-\tau)}DdW(\tau)\Big\|^2\Big\} \\ &\quad + 8\Big\|\frac{1}{\gamma^2}\int_0^t e^{H(t-\tau)}DD^T s(\tau)d\tau\Big\|^2 \\ &\triangleq 8\Xi_1(t) + 8\Xi_2(t) + 8\Xi_3(t) + 8\Xi_4(t).\end{aligned}$$

Then it suffices to show that $\limsup_{T\to\infty}(1/T)\int_0^T \Xi_j(t)dt < \infty$, $j = 1, 2, 3, 4$.

From Remark 4.3(ii), we have $\|e^{Ht}\bar{x}\|^2 \leq \rho^2 e^{-2\eta t}\|\bar{x}\|^2$; hence, we can show that $\limsup_{T\to\infty}(1/T)\int_0^T \Xi_1(t)dt = 0$. To prove the second term, since $\|s\|_\infty \leq \|g\|_\infty\|Q\|\rho/\eta =: M$, we have $\Xi_2^{1/2}(t) \leq \|B_i\|^2\|R^{-1}\|M\rho/\eta =: \bar{M}$, which leads to $\limsup_{T\to\infty}(1/T)\int_0^T \Xi_2(t)dt \leq \bar{M}^2$. Similarly, it can be shown that $\limsup_{T\to\infty}(1/T)\int_0^T \Xi_4(t)dt < \infty$. For the third part, note that by Itô isometry,

$$\Xi_3(t) = \mu\,\mathrm{Tr}\Big(\int_0^t e^{H(t-\tau)}DD^T e^{H^T(t-\tau)}d\tau\Big) \triangleq \mathrm{Tr}(Z(t)),$$

where $Z(0) = \mu DD^T$. Then, $Z(t)$ is a nondecreasing function, and converges to $Z \geq 0$ as $t \to \infty$, since H is Hurwitz. Hence, $\Xi_3(t) \leq \mathrm{Tr}(Z)$ for all $t \geq 0$ and $\limsup_{T\to\infty}(1/T)\int_0^T \Xi_3(t)dt \leq \mathrm{Tr}(Z)$. Similarly, we can show that $\mathbb{E}\{\|x(T)\|^2\} = o(T)$. This completes the proof. □

C.2 Proofs for Chapter 4

Here, we provide proofs for several of the results presented in Chapter 4.

Proof of Proposition 4.1. Part (i) is shown in [106]. For part (iv), the solution of (4.15) can be written as

$$s(t) = e^{-H^T t} s(0) + \int_0^t e^{-H^T (t-\sigma)} Q g(\sigma) d\sigma.$$

Then it is easy to show that with $s(0)$, $s(t)$ admits a unique solution in $\mathcal{C}_n^b$ [70, 71]. Part (iii) follows from [71, Theorems 3.1 and 4.1], since G and H are Hurwitz.

To prove parts (ii) and (v), let

$$\bar{J}_T(u, g) = \delta \log \mathbb{E}\big\{e^{\frac{1}{\delta}\bar{\phi}^1(x,g,u)}\big\},$$

where $\bar{\phi}^1(x, g, u) := \int_0^T \|x(t) - g(t)\|_Q^2 + \|u(t)\|_R^2 dt$ is defined in (4.12). We also define

$$\tilde{J}_T(u, g) = \delta \log \mathbb{E}\big\{e^{\frac{1}{\delta}\tilde{\phi}(x,g,u)}\big\},$$

where

$$\begin{aligned}\tilde{\phi}(x, g, u) &:= \bar{\phi}^1(x, g, u) + \|x(T)\|_P^2 + 2s^T(T)x(T) + \|s(T)\|_{P^{-1}}^2 \\ &= \bar{\phi}^1(x, g, u) + \|P^{1/2}x(T) + P^{-1/2}s(T)\|^2.\end{aligned}$$

Then clearly, $\bar{J}_T(u, g) \leq \tilde{J}_T(u, g)$, and

$$\limsup_{T\to\infty} \frac{1}{T}\bar{J}_T(u, g) = \bar{J}(u, g) \leq \limsup_{T\to\infty} \frac{1}{T}\tilde{J}_T(u, g).$$

Now, for any admissible controllers, it can be shown that

$$\bar{J}(u, g) \geq \limsup_{T\to\infty} \frac{1}{T}\int_0^T q(t)dt + \mu \operatorname{Tr}(PDD^T),$$

where $q(t) = \|g(t)\|_Q^2 - \|B^T s(t)\|_{R^{-1}}^2 + \gamma^{-2}\|D^T s(t)\|^2$.

On the other hand, by using Itô formula, (4.13), and (4.15) with the "completion of squares" method, we obtain

$$\begin{aligned}\tilde{J}_T(u, g) = \|\bar{x}\|_P^2 + 2x^T(0)s(0) + \mu T \operatorname{Tr}(PDD^T) + \|s(T)\|_{P^{-1}}^2 \\ + \int_0^T q(t)dt + \delta \log \mathbb{E}\{e^{\Upsilon_1(T)+\Upsilon_2(T)}\},\end{aligned}$$

where

$$\Upsilon_1(T) = \frac{1}{\delta}\int_0^T \|u(t) + R^{-1}(B^T Px(t) + B^T s(t))\|_R^2 dt$$
$$\Upsilon_2(T) = \varsigma \int_0^T (x^T(t)PD + s^T(t)D)dW(t) - \frac{\varsigma^2}{2}\int_0^T \|D^T Px(t) + D^T s(t)\|^2 dt,$$

where $\varsigma = 2\sqrt{\mu}/\delta$. Now, we introduce a change of probability measure [77]

$$d\bar{\mathbb{P}} = e^{\Upsilon_2(T)} d\mathbb{P},$$

where it can be verified that $e^{\Upsilon_2(t)}$ is a martingale on $[0, T]$ for any admissible controllers in $\mathcal{U}_{1,i}^d$ and $T \geq 0$ [114]. Then due to the Girsanov theorem [77, 114], $\bar{\mathbb{P}}$ is also a valid probability measure, and we can define the new expectation $\bar{\mathbb{E}}$ with respect to $\bar{\mathbb{P}}$. Therefore, since $s \in \mathcal{C}_n^b$, we have

$$\begin{aligned}\bar{J}(u, g) &\leq \limsup_{T\to\infty} \frac{1}{T}\tilde{J}_T(u, g) \\ &= \limsup_{T\to\infty}\Big[\frac{\delta}{T}\log \bar{\mathbb{E}}\{e^{\Upsilon_1(T)}\} + \frac{1}{T}\int_0^T q(t)dt\Big] + \mu\,\mathrm{Tr}(PDD^T).\end{aligned}$$

Then it is easy to see that the controller given in (4.14) is the corresponding optimal controller, and (4.17) is the minimum cost. This completes the proof. □

Proof of Theorem 4.2. For the first result of part (i), we consider the following relation:

$$\mathbb{E}\Bigg\{\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt\Bigg\} \leq 2\mathbb{E}\Bigg\{\int_0^T \Big\|f_N^*(t) - \frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t)\Big\|^2 dt\Bigg\} \quad \text{(C.1)}$$
$$+ 2T\sup_{t\geq 0}\Big\|\frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t) - g^*(t)\Big\|^2. \quad \text{(C.2)}$$

Let $e_i^*(t) = x_i^*(t) - \bar{x}_i^*(t)$ and $\Lambda_i(t) := \mathbb{E}\{e_i^*(t)(e_i^*(t))^T\}$ where $\bar{x}_i^*(t)$ is (4.19) with g^*. Then $e_i(0) = 0$ and $\Lambda_i(0) < \infty$ for all i. The SDE for $e_i^*(t)$ can be written as

$$de_i^* = G_i e_i^* dt + \sqrt{\mu} D_i dW_i(t). \tag{C.3}$$

Now, for (C.1), we have

$$\mathbb{E}\left\{\int_0^T \left\| \frac{1}{N}\sum_{i=1}^N e_i^*(t)\right\|^2 dt\right\} = \frac{1}{N^2}\operatorname{Tr}\Big(\int_0^T \sum_{i=1}^N \Lambda_i(t)dt\Big),$$

where the equality follows due to the fact that the Brownian motions are independent with all agents. It can be shown that $\Lambda_i(t)$ satisfies the following Lyapunov equation:

$$\frac{d\Lambda_i(t)}{dt} = G_i\Lambda_i(t) + \Lambda_i(t)G_i^T + \mu D_i D_i^T.$$

Note that G_i is Hurwitz for all agents. Then the above Lyapunov equation is monotonically nondecreasing, and converges to a positive definite matrix, say, $\Lambda_i > 0$, as $t \to \infty$ for all i. Since Θ is compact, there exists a positive definite matrix Λ independent of N such that

$$\mathbb{E}\left\{\int_0^T \left\| f_N^*(t) - \frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t)\right\|^2 dt\right\} \leq \frac{T}{N}\operatorname{Tr}(\Lambda).$$

Therefore (C.1) converges to zero as $N \to \infty$ for all $T \geq 0$. For (C.2), note that under Assumption 4.1,

$$\frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t) \equiv \int_{(\theta,x)\in\Theta\times X} \bar{x}_\theta^*(t)dF_N(\theta,x).$$

Then the convergence (C.2) follows from Lemma C.2(ii) in Appendix C.1, and Assumptions 4.1 and 4.2. The second statement of part (i) follows from the first one.

To prove the first statement of part (ii), consider

$$\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt \le 2\int_0^T \Big\|\frac{1}{N}\sum_{i=1}^N e_i^*(t)\Big\|^2 dt + 2T\sup_{t\ge 0}\Big\|\frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t) - g^*(t)\Big\|^2,$$

where $e_i^*(t)$ is defined in (C.3). Then the second term converges to zero as $N \to \infty$. Note that for each $t \ge 0$, $e_i^*(t)$ is a random vector with $\mathbb{E}\{e_i^*(t)\} = 0$ and $\mathbb{E}\{\|e_i^*(t)\|^2\} = \mathrm{Tr}(\Lambda_i(t)) \le \mathrm{Tr}(\Lambda_i) \le \mathrm{Tr}(\Lambda) < \infty$ for all i and $t \ge 0$. Hence, for each $t \ge 0$, $e_1^*(t), e_2^*(t), ...$ are mutually orthogonal random vectors. Also, it can be shown that by the integral test, for all $t \ge 0$,

$$\lim_{N\to\infty}\sum_{i=1}^N \frac{\mathrm{Tr}(\Lambda_i(t))\log^2(i)}{i^2} \le \lim_{N\to\infty}\sum_{i=1}^N \frac{\mathrm{Tr}(\Lambda)\log^2(i)}{i^2} < \infty.$$

Then $\lim_{N\to\infty}\|(1/N)\sum_{i=1}^N e_i^*(t)\| = 0$ a.s. for all $t \ge 0$ due to the law of large numbers [122, Theorem 5.2]. Since $\|(1/N)\sum_{i=1}^N e_i^*(t)\|$ is integrable on $[0, T]$ for all $T \ge 0$ [71, Lemma 5.2], $\lim_{N\to\infty}\int_0^T \|(1/N)\sum_{i=1}^N e_i^*(t)\|^2 dt = 0$ for all $T \ge 0$ a.s. which proves the first statement of part (ii).

For the second statement of part (ii), we have

$$\limsup_{T\to\infty}\frac{1}{T}\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt \le \limsup_{T\to\infty}\frac{2}{T}\int_0^T \Big\|\frac{1}{N}\sum_{i=1}^N e_i^*(t)\Big\|^2 dt + 2\sup_{t\ge 0}\Big\|\frac{1}{N}\sum_{i=1}^N \bar{x}_i^*(t) - g^*(t)\Big\|^2.$$

Then the second term converges to zero as $N \to \infty$. For the convergence of the first term, we follow the proof in [71, Lemma 5.3]. Let

$$e_i^*(t) = \sqrt{\mu}\int_{-\infty}^t e^{G_i(t-s)}D_i dW_i(s) - \sqrt{\mu}\int_{-\infty}^0 e^{G_i(t-s)}D_i dW_i(s) \triangleq w_i^1(t) + w_i^2(t).$$

It should be pointed out that the Brownian motion for $t < 0$ is just for mathematical purposes, since this process does not affect the entire process of $e_i^*(t)$ for all t and i [122]. Now, it suffices to show that (a.s.)

$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T}\int_0^T \Big\| \frac{1}{N}\sum_{i=1}^N w_i^1(t) \Big\|^2 dt = 0 \tag{C.4}$$

$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T}\int_0^T \Big\| \frac{1}{N}\sum_{i=1}^N w_i^2(t) \Big\|^2 dt = 0. \tag{C.5}$$

For (C.4), it can be verified that $m_N(t) := (1/N)\sum_{i=1}^N w_i^1(t)$ is a wide-sense stationary Gaussian process with mean-zero and the following auto-correlation function [122]:

$$R(\tau,t) = \mathbb{E}\{m_N(\tau)m_N^T(t)\} = \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\{w_i^1(\tau)w_i^{1,T}(t)\}.$$

Then from the result on estimation of the auto-correlation function in [71, Lemma 5.3] and [122], we have (a.s.)

$$\begin{aligned}&\limsup_{T\to\infty} \frac{1}{T}\int_0^T \|m_N(t)\|^2 dt \\ &= \frac{1}{N^2}\operatorname{Tr}(\mathbb{E}\{R(0,0)\}) \leq \frac{\mu\rho^2 \sup_{\theta\in\Theta}\operatorname{Tr}(D(\theta)D^T(\theta))}{2\eta N},\end{aligned}$$

which converges to zero as $N\to\infty$ with probability 1.

Let $z_i := \|\sqrt{\mu}\int_{-\infty}^0 e^{-G_i s}D_i dW_i(s)\|^2$. We can show that by using Itô formula, $\mathbb{E}\{z_i^2\} < \infty$ for all i, and hence $\lim_{N\to\infty}\sum_{i=1}^N \sup_{\theta\in\Theta}\mathbb{E}\{z_\theta^2\}/i^2 < \infty$. Then from the law of large numbers [122, Theorem 3.4], we can show that $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N (z_i - \mathbb{E}\{z_i\}) = 0$ almost surely. This implies that for any κ, there exists $N' := N(\kappa)$ such that for all $N \geq N'$, $\frac{1}{N}\sum_{i=1}^N z_i \leq \kappa$. Then with Remark 4.3(ii), for all $N \geq N'$, we have

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \Big\| \frac{1}{N}\sum_{i=1}^N w_i^2(t) \Big\|^2 dt \leq \limsup_{T\to\infty} \frac{1}{T}\int_0^T \kappa\rho^2 e^{-2\eta t} dt = 0, \text{ a.s.}$$

which establishes (C.5). This completes the proof of the theorem. □

Proof of Corollary 4.1. (i) From Theorem 4.2(ii), for each $c > 0$, there exists $N' := N(c)$ such that for all $N \geq N'$, we have $X(N,T) \leq c$ for all $T \geq 0$. Note that $X(N,T)$ is now a random variable that is bounded above by c with probability 1. Then, by using Hoeffding's inequality

[128], for any $s \geq 0$,

$$\mathbb{P}(X(N,T) - \mathbb{E}\{X(N,T)\} \geq s) \leq e^{-2s^2/c^2}.$$

The same bound can be obtained for $\mathbb{P}(X(N,T) - \mathbb{E}\{X(N,T)\} \leq -s)$. Then the result follows from their union bound.

(ii) For each $c(N)$ with $N \geq 1$, there exists $T' := T(c(N))$ such that for all $T \geq T'$, $(1/T)X(N,T) \leq c(N)$, which implies $X(N,T)$ is bounded above by $c(N)T$ with probability 1. Then, by using Hoeffding's inequality, for any $s \geq 0$, we get

$$\mathbb{P}(X(N,T) - \mathbb{E}\{X(N,T)\} \geq sT) \leq e^{-2s^2/c(N)^2}.$$

The remaining part is similar to part (i). This completes the proof.

□

Proof of Theorem 4.3. From Theorem 4.2, for each $N \geq 1$ and $\epsilon(N)$, there exists $T' := T(\epsilon(N))$ such that for all $T \geq T'$

$$\Big(\frac{1}{T}\int_0^T \|f_N^*(t) - g^*(t)\|^2 dt\Big)^{1/2} \leq \epsilon(N), \text{ a.s.}$$

Note that in view of Theorem 4.2(ii) and Corollary 4.1, for large N, we can choose a small enough $\epsilon(N)$ to arrive at the above relation.

Due to Proposition 4.1(iii), there exist M_1 and T_1 such that

$$\int_0^T \|x_i^*(t)\|^2 dt \leq M_1 T, \ \forall T \geq T_1, \text{ a.s.} \tag{C.6}$$

Moreover, note that due to Proposition 4.2, $J_{1,i}^N(u_i^*, u_{-i}^*) \leq C$ where $C \geq 0$ is independent on N. Since $\mathcal{U}_{1,i}^d \subseteq \mathcal{U}_{1,i}^c$, which implies $\inf_{u_i \in \mathcal{U}_{1,i}^c} J_{1,i}^N(u_i, u_{-i}^*) \leq J_{1,i}^N(u_i^*, u_{-i}^*)$, we may consider $u_i \in \mathcal{U}_{1,i}^c$ with the property that there exist M_2 and T_2 such that

$$\int_0^T \|x_i(t)\|^2 dt \leq M_2 T, \ \forall T \geq T_2, \text{ a.s.} \tag{C.7}$$

Note that in (C.6) and (C.7), M_1 and M_2 do not depend on N due to the compactness of Θ. Furthermore, since $g^* \in \mathcal{C}_n^b$, from Proposition 4.1, there exist M_3 and T_3 such that

$$\int_0^T \|g^*(t)\|^2 dt \le M_3 T, \ \forall T \ge T_3.$$

Let $\bar{T} := \max\{T', T_1, T_2, T_3\}$, $\bar{M}_1 := 4\max\{M_1, M_2\}$, $\bar{M}_2 := 4\max\{M_1, M_3\}$, and $\bar{M}_3 := 4\max\{M_2, M_3\}$.

Now, from Lemma C.1(i), for all $T \ge \bar{T}$, we have (a.s.)

$$\begin{aligned}
&\int_0^T \|x_i^* - f_N^*\|_Q^2 + \|u_i^*\|_R^2 dt \\
&\le \int_0^T \|x_i^* - g^*\|_Q^2 + \|u_i^*\|_R^2 dt + \|Q\|\epsilon(N)^2 T + 2\|Q\|(\bar{M}_2 T)^{1/2}\epsilon(N)T^{1/2}.
\end{aligned}$$

Since exponentiating, taking expectation and logarithm, and taking the limit of the expressions on both sides do not change the direction of the inequality, it follows that

$$\begin{aligned}
&\limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\phi_i^1(x^*, f_N^*, u^*)}\} \\
&\le \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\bar{\phi}_i^1(x^*, g^*, u^*)}\} + 2\|Q\|\epsilon(N)^2 + 2\|Q\|\bar{M}_2^{1/2}\epsilon(N), \qquad \text{(C.8)}
\end{aligned}$$

where (C.8) is the performance index of (4.11) for agent i when agent i uses the optimal decentralized controller in (4.21).

Furthermore, from Lemma C.1(ii), for all $T \ge \bar{T}$,

$$\begin{aligned}
&\int_0^T \|x_i - g^*\|_Q^2 + \|u_i\|_R^2 dt \\
&\le \int_0^T \|x_i - f_N^{-i^*}\|_Q^2 + \|u_i\|_R^2 dt + 2\|Q\|(\bar{M}_3 T)^{1/2}\epsilon(N)T^{1/2} \\
&\quad + \frac{2\|Q\|}{N}(\bar{M}_3 T)^{1/2}(\bar{M}_1 T)^{1/2}.
\end{aligned}$$

Again, since exponentiating, taking expectation and logarithm, and taking the limit do not change the direction of the inequality, it follows that

$$\begin{aligned}
&\limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\bar{\phi}_i^1(x^*,g^*,u^*)}\} \\
&\leq \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\bar{\phi}_i^1(x,g^*,u)}\} \\
&\leq \limsup_{T\to\infty} \frac{\delta}{T} \log \mathbb{E}\{e^{\frac{1}{\delta}\phi_i^1(x,f_N^{-i^*},u)}\} + 2\|Q\|\bar{M}_3^{1/2}\epsilon(N) + \frac{2\|Q\|}{N}\bar{M}_3^{1/2}\bar{M}_1^{1/2},
\end{aligned}$$

where $\bar{\phi}_i^1(x, g^*, u)$ is the modified version of (4.12) when agent i is under the centralized controller $u_i \in \mathcal{U}_{1,i}^c$. Note that the first inequality follows from the fact that the optimal decentralized controller in (4.21) solves the robust tracking problem in Proposition 4.1.

Therefore, we have

$$\begin{aligned}
J_{1,i}^N(u_i^*, u_{-i}^*) \leq \inf_{u_i\in\mathcal{U}_{1,i}^c} J_{1,i}^N(u_i, u_{-i}^*) + 2\|Q\|\epsilon(N)^2 + 2\|Q\|\bar{M}_2^{1/2}\epsilon(N) \\
+ 2\|Q\|\bar{M}_3^{1/2}\epsilon(N) + \frac{2\|Q\|}{N}\bar{M}_3^{1/2}\bar{M}_1^{1/2},
\end{aligned}$$

which implies

$$J_{1,i}^N(u_i^*, u_{-i}^*) \leq \inf_{u_i\in\mathcal{U}_{1,i}^c} J_{1,i}^N(u_i, u_{-i}^*) + O(\epsilon(N)) + O(\frac{1}{N}).$$

Note that $\epsilon(N)$ is a constant that is dependent on N, which converges to zero as $N \to \infty$ due to the previous argument and Theorem 4.2. This completes the proof of the theorem. □

Proof of Proposition 4.4. We need to prove parts (ii) and (v). Part (iii) was shown in Lemma C.3 in Appendix C.1. The proofs of the remaining parts are similar to that of Proposition 4.1.

By using the Itô formula, (4.13), and (4.15), it can be shown that

$$\begin{aligned}
\mathbb{E}\int_0^T \|x(t)\|_Q^2 dt = \mathbb{E}\{\|x(0)\|_P^2 - \|x(T)\|_P^2\} + \mu T \operatorname{Tr}(PDD^T) \\
+ \mathbb{E}\int_0^T x^T(t)(PBR^{-1}B^TP - \frac{1}{\gamma^2}PDD^TP)x(t)dt \\
+ 2\mathbb{E}\int_0^T (u^T(t)B^TP + v^T(t)D^TP)x(t)dt,
\end{aligned}$$

and

$$-2\mathbb{E}\int_0^T h^T(t)Qx(t)dt = 2\mathbb{E}\{s(0)^T x(0) - s^T(T)x(T)\}$$
$$+ 2\mathbb{E}\int_0^T s^T(t)(BR^{-1}B^TP - \frac{1}{\gamma^2}DD^TP)x(t)dt$$
$$+ 2\mathbb{E}\int_0^T s^T(t)(Bu(t) + Dv(t))dt.$$

From the "completion of squares" method, we have

$$\mathbb{E}\int_0^T \|x(t) - h(t)\|_Q^2 + \|u(t)\|_R^2 - \gamma^2\|v(t)\|^2 dt$$
$$= \mathbb{E}\{\|x(0)\|_P^2 - \|x(T)\|_P^2 + 2s^T(0)x(0) - 2s^T(T)x(T)\} + \mu T\,\mathrm{Tr}(PDD^T)$$
$$+ \int_0^T q(t)dt + \mathbb{E}\int_0^T \|u(t) + R^{-1}B^TPx(t) + R^{-1}B^Ts(t)\|_R^2 dt$$
$$- \mathbb{E}\int_0^T \gamma^2\|v(t) - \frac{1}{\gamma^2}D^TPx(t) - \frac{1}{\gamma^2}D^Ts(t)\|^2 dt,$$

where $q(t) = \|h(t)\|_Q^2 - \|B^Ts(t)\|_{R^{-1}}^2 + \gamma^{-2}\|D^Ts(t)\|^2$. Then due to the fact that $s \in \mathcal{C}_n^b$ and $\mathbb{E}\{\|x(T)\|^2\} = o(T)$ shown in Lemma C.3 in Appendix C.1, we have

$$\bar{J}(u,v,h) = \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\int_0^T \|u(t) + R^{-1}B^T(Px(t) + s(t))\|_R^2 dt$$
$$- \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\int_0^T \gamma^2\|v(t) - \frac{1}{\gamma^2}D^T(Px(t) + s(t))\|^2 dt$$
$$+ \limsup_{T\to\infty}\frac{1}{T}\int_0^T q(t)dt + \mu\,\mathrm{Tr}(PDD^T).$$

Since H is Hurwitz, (4.25) and (4.26) constitute the saddle-point equilibrium of the infinite-horizon LQ stochastic zero-sum differential game [61], and the corresponding saddle-point value is (4.17) with $g(t)$ replaced by $h(t)$. This completes the proof. □

Proof of Theorem 4.6. The proof is similar to that of Theorem 4.3. We define

$$\epsilon(N) := \Big(\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big\{\int_0^T \|f_N^*(t) - h^*(t)\|^2 dt\Big\}\Big)^{\frac{1}{2}}.$$

Note that $\epsilon(N)$ defined above is similar to that in Theorem 4.3, since by

Theorem 4.5(i) and Corollary 4.2, $\epsilon(N)$ can be taken to be arbitrarily small when N is arbitrarily large.

From Lemma C.3 in Appendix C.1, there exists M_1 such that

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big\{\int_0^T \|x_i^*(t)\|^2 dt\Big\} \leq M_1.$$

Moreover, note that due to Proposition 4.5, $J_{2,i}^N(u_i^*, u_{-i}^*) \leq C$ where $C \geq 0$ is independent of N. Since $\mathcal{U}_{2,i}^d \subseteq \mathcal{U}_{2,i}^c$, which implies $\inf_{u_i \in \mathcal{U}_{2,i}^c} J_{2,i}^N(u_i, u_{-i}^*) \leq J_{2,i}^N(u_i^*, u_{-i}^*)$, we may consider $u_i \in \mathcal{U}_{2,i}^c$ with the property that with the worst-case disturbance in (4.30), there exists M_2 such that

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big\{\int_0^T \|x_i(t)\|^2 dt\Big\} \leq M_2.$$

Note that due to the compactness of Θ, M_1 and M_2 are not dependent on N. We also use M_3 as defined in the proof of Theorem 4.3. Then we can define $\bar{M}_1$, $\bar{M}_2$ and $\bar{M}_3$ in a similar way to that in the proof of Theorem 4.3.

We introduce the performance index of the stochastic zero-sum differential game for agent i:

$$\bar{J}_{2,i}(u_i^*, v_i^*, h^*) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\big\{\bar{\phi}_i^2(x^*, h^*, u^*, v^*)\big\},$$

where $\bar{\phi}_i^2(x^*, h^*, u^*, v^*)$ is (4.24) for agent i when $u_i \equiv u_i^*$ and $v_i \equiv v_i^*$. Note that due to the definition of saddle-point equilibrium [61], $\bar{J}_{2,i}(u_i^*, v_i^*, h^*) = \inf_{u_i \in \mathcal{U}_{2,i}^d} \bar{J}_{2,i}(u_i, v_i^*, h^*) = \sup_{v_i \in \mathcal{V}_i} \bar{J}_{2,i}(u_i^*, v_i, h^*)$.

Now, by using Lemma C.1(i) and the above equality condition, we have

$$\begin{aligned} J_{2,i}^N(u_i^*, u_{-i}^*) &= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\phi_i^2(x^*, f_N^*, u^*, v^*)\} \\ &\leq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\big\{\bar{\phi}_i^2(x^*, h^*, u^*, v^*)\big\} + \|Q\|\epsilon(N)^2 + 2\|Q\|\bar{M}_2^{1/2}\epsilon(N) \\ &= \bar{J}_{2,i}(u_i^*, v_i^*, h^*) + O(\epsilon(N)). \end{aligned} \tag{C.9}$$

Also, by using Lemma C.1(ii) and the above equality condition,

$$
\begin{aligned}
\bar{J}_{2,i}(u_i^*, v_i^*, h^*) &\leq \bar{J}_{2,i}(u_i, v_i^*, h^*) \\
&\leq \sup_{v_i \in \mathcal{V}_i} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\{\phi_i^2(x, f_N^{-i^*}, u, v)\} \\
&\quad + 2\|Q\|\bar{M}_3^{1/2}\epsilon(N) + \frac{2\|Q\|(\bar{M}_3\bar{M}_1)^{\frac{1}{2}}}{N} \\
&= J_{2,i}^N(u_i, u_{-i}^*) + O(\epsilon(N)) + O(1/N). \qquad \text{(C.10)}
\end{aligned}
$$

Then from (C.9) and (C.10), we have the desired result. This completes the proof of the theorem. □

APPENDIX D

APPENDIX FOR CHAPTER 5

D.1 Block Matrices for Nonsymmetric RDEs

From Assumption 5.1(c), $\pi = (\pi_1, ..., \pi_K)$. The following block matrices are used in RDEs in (5.17) and (5.29):

$$
\begin{aligned}
\mathbb{A} &= \text{diag}\{A_1, ..., A_K\} \\
\mathbb{B}_1 &= \pi \otimes -B_0 R_0^{-1}(L^T R^{-1} B^T - F^T) \\
\mathbb{B}_2 &= I_K \otimes (BR^{-1}B^T),\ \mathbb{B}_3 = \mathbf{1}_K \otimes (-(BR^{-1}L - F)R_0^{-1}B_0^T) \\
\mathbb{B}_4 &= \pi \otimes (BR^{-1}L - F)R_0^{-1}(L^T R^{-1} B^T - F^T),\ \mathbb{B}_K = \mathbf{1}_K \otimes \mathbb{B}_4 \\
\mathbb{Q}_1 &= \pi \otimes -Q_0 H_0,\ \mathbb{Q}_2 = \mathbf{1}_K \otimes (-H_0^T Q_0) \\
\mathbb{Q}_3 &= \begin{pmatrix} Q - QH\pi_1 & -QH\pi_2 & \dots & -QH\pi_K \\ -QH\pi_1 & Q - QH\pi_2 & \dots & -QH\pi_K \\ \vdots & \vdots & \ddots & \vdots \\ -QH\pi_1 & -QH\pi_2 & \dots & Q - QH\pi_K \end{pmatrix} \\
\mathbb{Q}_4 &= \pi \otimes H_0^T Q H_0,\ \mathbb{Q}_K = \mathbf{1}_K \otimes \mathbb{Q}_4 \\
\bar{q}(t) &= \begin{pmatrix} q_1(t) & \dots & q_K(t) \end{pmatrix}.
\end{aligned}
$$

D.2 Auxiliary Lemma

Lemma D.1. *Let the pair $(\Psi(t), \Pi(t))$ be the solution of the following differential equation:*

$$
\frac{d}{dt}\begin{pmatrix} \Psi(t) \\ \Pi(t) \end{pmatrix} = \begin{pmatrix} \mathcal{A}_1 & -\mathcal{B}_1 \\ -\mathcal{A}_2 & \mathcal{B}_2 \end{pmatrix} \begin{pmatrix} \Psi(t) \\ \Pi(t) \end{pmatrix},\ \begin{pmatrix} \Psi(T) \\ \Pi(T) \end{pmatrix} = \begin{pmatrix} I \\ 0 \end{pmatrix},
$$

where $\Psi(t)$ is invertible for all $t \in [0, T]$. Then the RDE in (5.29) has a unique solution $\Lambda(t) = \Pi(t)\Psi^{-1}(t)$ for all $t \in [0, T]$.

Proof. Note that

$$\frac{d\Psi^{-1}(t)}{dt} = -\Psi^{-1}(t)\frac{d\Psi(t)}{dt}\Psi^{-1}(t) = -\Psi^{-1}(t)\mathcal{A}_1 + \Psi^{-1}(t)\mathcal{B}_1\Pi(t)\Psi^{-1}(t),$$

which implies

$$\begin{aligned}\frac{d\Lambda(t)}{dt} &= (-\mathcal{A}_2\Psi(t) + \mathcal{B}_2\Pi(t))\Psi^{-1}(t) \\ &\quad + \Pi(t)(-\Psi^{-1}(t)\mathcal{A}_1 + \Psi^{-1}(t)\mathcal{B}_1\Pi(t)\Psi^{-1}(t)).\end{aligned}$$

This proves the lemma. □

D.3 Proof of Lemma 5.2

The existence and uniqueness of the optimal controller follows from a similar argument as discussed in [120, 121, 119].

To prove the optimality of the controller given in (5.26), we first consider the following variations of SDEs with $\delta u_0 \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^p)$ and $k \in \mathcal{K}$:

$$\begin{aligned}d\delta x_0(t) &= A_0\delta x_0(t)dt + B_0\delta u_0(t)dt \\ d\delta z_k(t) &= A_k\delta z_k(t)dt + BR^{-1}B^T\delta\bar{p}_k(t)dt - (BR^{-1}L - F)\delta u_0(t)dt \\ \delta\bar{p}_k(t) &= -A_k^T\delta\bar{p}_k(t)dt + Q(\delta z_k(t) - H\delta z(t))dt,\end{aligned}$$

where it can be seen that $\delta x_0(0) = 0$, $\delta z_k(0) = 0$ and $\delta\bar{p}_k(T) = 0$. We also note that $\delta z(t) = \sum_{k=1}^K \pi_k\delta z_k(t)$. Since the cost functional is convex in x_0 and strictly convex in u_0, u_0^* is the corresponding optimal controller if and only if it satisfies the first-order condition

$$\begin{aligned}0 = \frac{\delta\bar{J}(\bar{u}_0)}{2} = \mathbb{E}\int_0^T \Big[&\delta x_0^T(t)Q(x_0(t) - H_0 z(t)) \\ &+ \sum_{k=1}^K \pi_k\delta z_k^T(t)H_0^T Q_0(H_0 z(t) - x_0(t)) + \delta u_0^T(t)R_0 u_0(t)\Big]dt.\end{aligned} \tag{D.1}$$

Now, by Itô formula,

$$
\begin{aligned}
d\delta x_0^T(t)p_0(t) = {} & \delta u_0^T(t)B_0^T p_0(t)dt + \delta z^T(t) + \delta x_0^T(t)Q(x_0(t) - H_0 z(t))dt \\
& + \delta x_0^T(t)q_0(t)dW_0(t),
\end{aligned}
$$

which together with boundary conditions implies

$$
0 = \mathbb{E}\int_0^T \Big[\delta u_0^T(t)B_0^T p_0(t) + \delta x_0^T(t)Q(x_0(t) - H_0 z(t))\Big]dt. \tag{D.2}
$$

By Itô formula,

$$
\begin{aligned}
d\delta z_k^T(t)\lambda_k(t) = {} & \Big[\delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t) - \delta u_0^T(t)(L^T R^{-1}B^T - F^T)\lambda_k(t) \\
& + \delta z_k^T(t)H_0^T Q_0(H_0 z(t) - x_0(t)) - \delta z_k^T(t)Q\xi_k(t) \\
& + \delta z_k^T(t)H^T Q\sum_{k=1}^K \pi_k\xi_k(t)\Big]dt + \delta z^T(t)q_k(t)dW_0(t),
\end{aligned}
$$

which implies

$$
\begin{aligned}
0 = \mathbb{E}\int_0^T & \Big[\delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t) - \delta u_0^T(t)(L^T R^{-1}B^T - F^T)\lambda_k(t) \\
& + \delta z_k^T(t)H_0^T Q_0(H_0 z(t) - x_0(t)) - \delta z_k^T(t)Q\xi_k(t) \\
& + \delta z_k^T(t)H^T Q\sum_{k=1}^K \pi_k\xi_k(t)\Big]dt.
\end{aligned}
$$

Multiplying both sides by π_k and summing over K yield

$$
\begin{aligned}
0 = \mathbb{E}\int_0^T & \Big[\sum_{k=1}^K \pi_k\delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t) - \sum_{k=1}^K \pi_k\delta u_0^T(t)(L^T R^{-1}B^T - F^T)\lambda_k(t) \\
& - \sum_{k=1}^K \pi_k\delta z_k^T(t)Q\xi_k(t) + \sum_{k=1}^K \pi_k\delta z_k^T(t)H_0^T Q_0(H_0 z(t) - x_0(t)) \\
& + \sum_{k=1}^K \pi_k\delta z_k^T(t)H^T Q\sum_{k=1}^K \pi_k\xi_k(t)\Big]dt.
\end{aligned} \tag{D.3}
$$

By Itô formula,

$$
d\delta\bar{p}_k^T(t)\xi_k(t) = \Big[\delta z_k^T(t)Q\xi_k(t) - \delta z^T(t)H^T Q\xi_k(t) - \delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t)\Big]dt,
$$

which implies

$$0=\mathbb{E}\int_0^T\Big[\delta z_k^T(t)Q\xi_k(t)-\delta z^T(t)H^TQ\xi_k(t)-\delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t)\Big]dt.$$

Multiplying both sides by π_k and summing over K yield

$$\begin{aligned}0=\mathbb{E}\int_0^T\Big[&\sum_{k=1}^K\pi_k\delta z_k^T(t)Q\xi_k(t)-\delta z^T(t)H^TQ\sum_{k=1}^K\pi_k\xi_k(t)\\&-\sum_{k=1}^K\pi_k\delta\bar{p}_k^T(t)BR^{-1}B^T\lambda_k(t)\Big]dt.\end{aligned}\tag{D.4}$$

Then from (D.1)-(D.4), we have

$$0=\mathbb{E}\int_0^T\delta u^T(t)\Big[R_0u_0(t)-B_0^Tp_0(t)+\sum_{k=1}^K\pi_k(L^TR^{-1}B^T-F^T)\lambda_k(t)\Big]dt.$$

This leads to the desired result, and completes the proof.

REFERENCES

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

[2] G. Nair, F. Fagnani, S. Zampieri, and R. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[3] R. Murray, K. Åstrom, S. Boyd, R. Brockett, and G. Stein, "Future directions in control in an information-rich world," *IEEE Control Systems Magazine*, vol. 23, no. 2, pp. 20–33, Apr. 2003.

[4] K. D. Kim and P. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proc. of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, May 2012.

[5] X. Li and V. Jilkov, "Survey of maneuvering target tracking. Part V: Multiple-model methods," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, Oct 2005.

[6] T. Fortmann, Y. Bar-Shalom, M. Scheffe, and S. Gelfand, "Detection thresholds for tracking in clutter–a connection between estimation and signal processing," *IEEE Transactions on Automatic Control*, vol. 30, no. 3, pp. 221–229, Mar 1985.

[7] A. C. Antoulas and D. C. Sorensen, "Approximation of large-scale dynamical systems: An overview," *Int. J, Appl. Comput. Sci.*, vol. 11, no. 5, pp. 1093–1121, 2001.

[8] C. Zhang, "Centralized and decentrlized control with limited information," Ph.D. dissertation, University of Illinois at Urbana and Champaign, 2008.

[9] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, vol. 36, no. 9, pp. 1249 – 1274, 2000.

[10] W. Saad, Z. Han, H. Poor, and T. Başar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, Sept 2012.

[11] G. Malewicz, M. H. Austern, A. J. C. Bik, J. C. Dehnert, and I. Horn, "Pregel: A system for large-scale graph processing," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, 2010, pp. 135–146.

[12] K. Guruharsha, J. F. Rual, B. Zhai, J. Mintseris, P. Vaidya, N. Vaidya, C. Beekman, C. Wong, D. Rhee, O. Cenaj, E. McKillip, S. Shah, M. Stapleton, K. Wan, C. Yu, B. Parsa, J. Carlson, X. Chen, B. Kapadia, K. VijayRaghavan, S. Gygi, S. E. Celniker, R. Obar, and S. Artavanis-Tsakonas, "A protein complex network of drosophila melanogaster," *Cell*, vol. 147, no. 3, pp. 690–703, 2011.

[13] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[14] M. Athans, R. Ku, and S. Gershwin, "The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty," *IEEE Transactions on Automatic Control*, vol. 22, no. 3, pp. 491–495, Jun 1977.

[15] R. Ku and M. Athans, "Further results on the uncertainty threshold principle," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 866 – 868, Oct. 1977.

[16] W. D. Koning, "Infinite horizon optimal control of linear discrete time systems with stochastic parameters," *Automatica*, vol. 18, no. 4, pp. 443 – 453, 1982.

[17] H. J. Chizeck and Y. Ji, "Optimal quadratic control of jump linear systems with Gaussian noise in discrete-time," in *Proceedings of the 27th Conference on Decision and Control*, Austin, Texas, 1998, pp. 1989–1993.

[18] O. Costa, M. Fragoso, and M. Todorov, *Continuous-Time Markov Jump Linear Systems*. Springer, 2013.

[19] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

[20] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Trans. on Autom. Control*, vol. 19, pp. 494–500, Oct. 1974.

[21] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[22] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. on Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004.

[23] E. Garone, B. Sinopoli, A. Goldsmith, and A. Casavola, "LQG control for MIMO systems over multiple erasure channels with perfect acknowledgment," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 450–456, Feb. 2012.

[24] E. Garone, B. Sinopoli, and A. Casavola, "LQG control over lossy TCP-like networks with probabilistic packet acknowledgement," *International Journal of Systems, Control, and Communications*, vol. 2, no. 1/2/3, pp. 55–81, 2010.

[25] M. Trivellato and N. Benvenuto, "State control in networked control systems under packet drops and limited transmission bandwidth," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 611–622, 2010.

[26] C.-C. Chang and S. Lall, "An explicit solution for optimal two-player decentralized control over TCP erasure channels with state feedback," in *Proceddings of the American Control Conference*, June 2011, pp. 4717–4722.

[27] Y. Mo, E. Garone, and B. Sinopoli, "LQG control with Markovian packet loss," in *European Control Conference, 2013*, July 2013, pp. 2380–2385.

[28] N. Elia, "Remote stabilization over fading channels," *Systems and Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.

[29] N. Elia and J. Eisenbeis, "Limitations of linear control over packet drop networks," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 826–841, April 2011.

[30] E. I. Silva and S. A. Pulgar, "Control of LTI plants over erasure channels," *Automatica*, vol. 47, no. 8, pp. 1729–1736, 2011.

[31] P. Seiler and R. Sengupta, "An H^∞ approach to networked control," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 356–364, March 2005.

[32] J. Geromel, A. Goncalves, and A. Fioravanti, “Dynamic output feedback control of discrete-time Markov jump linear systems through linear matrix inequalities,” *SIAM Journal on Control and Optimization*, vol. 48, no. 2, pp. 573–593, 2009.

[33] M. Sahebsara, T. Chen, and S. L. Shah, “Optimal H_∞ filtering in networked control systems with multiple packet dropouts,” *Systems and Control Letters*, vol. 57, no. 9, pp. 696–702, 2008.

[34] H. Ishii, “H^∞ control with limited communication and message losses,” *Systems and Control Letters*, vol. 57, no. 4, pp. 322–331, 2008.

[35] Z. Pan and T. Başar, “H^∞ control of Markovian jump systems and solutions to associated piecewise-deterministic differential games.” *In G.J. Olsder, editor, Annals of Dynamic Games*, vol. 2, pp. 61–94, 1995, Birkhäuser.

[36] O. Costa, M. Fragoso, and R. Marques, *Discrete-Time Markov Jump Linear Systems*. Springer-Verlag, 2005.

[37] J. Lee and G. Dullerud, “Optimal disturbance attenuation for discrete-time switched and Markovian jump linear systems,” *SIAM Journal on Control and Optimization*, vol. 45, no. 4, pp. 1329–1358, 2006.

[38] Z. Wang, F. Yang, D. W. C. Ho, and X. Liu, “Robust H^∞ control for networked systems with random packet losses,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 4, pp. 916–924, 2007.

[39] Z. Wang, D. W. Ho, Y. Liu, and X. Liu, “Robust control for a class of nonlinear discrete time-delay stochastic systems with missing measurements,” *Automatica*, vol. 45, no. 3, pp. 684 – 691, 2009.

[40] D. Wang, J. Wang, and W. Wang, “H_∞ controller design of networked control systems with Markov packet dropouts,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 3, pp. 689–697, May 2013.

[41] Y. Shoukry, J. Araujo, P. Tabuada, M. Srivastava, and K. H. Johansson, “Minimax control for cyber-physical systems under network packet scheduling attacks,” in *Proceedings of the 2nd ACM international conference on High confidence networked systems*, New York, NY, USA, 2013, pp. 93–100.

[42] N. Nahi, “Optimal recursive estimation with uncertain observation,” *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 457–462, Jul 1969.

[43] O. L. V. Costa and S. Guerra, "Stationary filter for linear minimum mean square error estimator of discrete-time markovian jump systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1351–1356, 2002.

[44] Y. Mo and B. Sinopoli, "A characterization of the critical value for Kalman filtering with intermittent observations," in *Proc. 47th Conf. on Decision and Control*, Dec. 2008, pp. 2692–2697.

[45] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," *IEEE Trans. on Autom. Control*, vol. 54, no. 2, pp. 386–390, Feb. 2009.

[46] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Critical value for second order system," in *Proc. of 18th IFAC World Congress*, 2011.

[47] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598 – 607, 2007.

[48] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. 43rd IEEE Conf. on Decision and Control*, Dec. 2004, pp. 4180–4186.

[49] S. Kluge, K. Reif, and M. Brokate, "Stochastic stability of the extended Kalman filter with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 514–518, 2010.

[50] L. Shi, M. Epstein, and R. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. on Autom. Control*, vol. 55, no. 3, pp. 594–604, March 2010.

[51] M. Epstein, L. Shi, A. Tiwari, and R. M. Murray, "Probabilistic performance of state estimation across a lossy network," *Automatica*, vol. 44, no. 12, pp. 3046 – 3053, 2008.

[52] A. Censi, "Kalman filtering with intermittent observations: Convergence for semi-markov chains and an intrinsic performance measure," *IEEE Trans. on Autom. Control*, vol. 56, no. 2, pp. 376–381, 2011.

[53] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Tail distribution and critical value," *IEEE Trans. on Autom. Control*, vol. 57, no. 3, pp. 677–689, 2012.

[54] S. Kar, B. Sinopoli, and J. M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *IEEE Trans. on Autom. Control*, vol. 57, no. 2, pp. 405 – 420, Feb. 2012.

[55] C. E. De Souza and M. Fragoso, "H^{∞} filtering for discrete-time linear systems with Markovian jumping parameters," in *Proceedings of the 36th IEEE Conference on Decision and Control (CDC)*, 1997, pp. 2181–2186.

[56] Z. Wang, F. Yang, D. Ho, and X. Liu, "Robust H_{∞} filtering for stochastic time-delay systems with missing measurements," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2579–2587, July 2006.

[57] H. Gao and T. Chen, "H^{∞} estimation for uncertain systems with limited communication capacity," *IEEE Transactions on Automatic Control*, vol. 52, no. 11, pp. 2070–2084, 2007.

[58] A. Goncalves, A. Fioravanti, and J. Geromel, "H^{∞} filtering of discrete-time Markov jump linear systems through linear matrix inequalities," *IEEE Trans. on Autom. Control*, vol. 54, no. 6, pp. 1347–1351, 2009.

[59] Z. Wang, H. Dong, B. Shen, and H. Gao, "Finite-horizon H_{∞} filtering with missing measurements and quantization effects," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1707–1718, 2013.

[60] D. Fudenberg and J. Tirole, *Game Theory*. The MIT Press, 1991.

[61] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. SIAM, 1999.

[62] H. Yin, P. Mehta, S. Meyn, and U. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 920–935, 2012.

[63] Z. Ma, D. Callaway, and I. Hiskens, "Decentralized charging control of large populations of plug-in electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 67–78, Jan 2013.

[64] R. Couillet, S. Perlaza, H. Tembine, and M. Debbah, "Electrical vehicles in the smart grid: A mean field game analysis," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1086–1096, July 2012.

[65] A. C. Kizilkale and R. P. Malhame, "Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads," in *Proceedings of the 19th IFAC World Congress*, Cape Town, South Africa, Aug. 2014, pp. 1867–1872.

[66] J. M. Lasry and P. L. Lions, "Mean field games," *Jap. J. Math.*, vol. 2, no. 1, pp. 229–260, 2007.

[67] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, "Markov perfect industry dynamics with many firms," *Econometrica*, vol. 76, no. 6, pp. 1375–1411, 2008.

[68] M. Huang, P. Caines, and R. Malhame, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions," in *Proceedings of the 42nd IEEE CDC*, Dec. 2003, pp. 98–103.

[69] Q. Zhu, H. Tembine, and T. Başar, "Hybrid risk-sensitive mean-field stochastic differential games with application to molecular biology," in *Proceedings of the 50th IEEE CDC and ECC*, Dec 2011, pp. 4491–4497.

[70] M. Huang, P. Caines, and R. Malhame, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

[71] T. Li and J.-F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1643–1660, 2008.

[72] B. Wang and J. Zhang, "Mean field games for large-population multiagent systems with Markov jump parameters," *SIAM Journal on Control and Optimization*, vol. 50, no. 4, pp. 2308–2334, 2012.

[73] H. Tembine, Q. Zhu, and T. Başar, "Risk-sensitive mean field games," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 835–850, 2014.

[74] H. Tembine, Q. Zhu, and T. Başar, "Risk-sensitive mean-field stochastic differential games," in *Proceedings of the 18th IFAC World Congress*, Milan, Italy, Aug. 2011, pp. 3222–3227.

[75] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, Apr. 1973.

[76] P. Whittle, *Risk-sensitive Optimal Control*. John Wiley & Sons, 1990.

[77] W. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. Springer, 2006.

[78] T. Runolfsson, "The equivalence between infinite-horizon optimal control of stochastic systems with exponential-of-integral performance index and stochastic differential games," *IEEE Transactions on Automatic Control*, vol. 39, no. 8, pp. 1551–1563, Aug. 1994.

[79] W. Fleming and W. McEneaney, “Risk-sensitive control on an infinite time horizon,” *SIAM Journal on Control and Optimization*, vol. 33, no. 6, pp. 1881–1915, 1995.

[80] Z. Pan and T. Başar, “Model simplification and optimal control of stochastic singularly perturbed systems under exponentiated quadratic cost,” *SIAM Journal on Control and Optimization*, vol. 34, no. 5, pp. 1734–1766, 1996.

[81] H. Nagai, “Bellman equations of risk-sensitive control,” *SIAM Journal on Control and Optimization*, vol. 34, no. 1, pp. 74–101, 1996.

[82] T. Duncan, “Linear-exponential-quadratic Gaussian control,” *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2910–2911, Nov 2013.

[83] T. Başar, “Nash equilibria of risk-sensitive nonlinear stochastic differential games,” *Journal of Optimization Theory and Applications*, vol. 100, no. 3, pp. 479–498, 1999.

[84] M. Huang, “Large-population LQG games involving a major player: The Nash certainty equivalence principle,” *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.

[85] S. L. Nguyen and M. Huang, “Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players,” *SIAM Journal on Control and Optimization*, vol. 50, no. 5, pp. 2907–2937, 2012.

[86] M. Nourian and P. Caines, “ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents,” *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, 2013.

[87] H. Von Stackelberg, *The Theory of Market Economy.* Oxford University Press, 1952.

[88] G. Papavassilopoulos and J. B. Cruz, “Nonclassical control problems and Stackelberg games,” *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 155–166, Apr 1979.

[89] M. Simaan and J. Cruz, “A Stackelberg solution for games with many players,” *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 322–324, Jun. 1973.

[90] T. Başar and H. Selbuz, “Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems,” *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 166–179, Apr 1979.

[91] G. Freiling, G. Jank, and R. Lee, "Existence and uniqueness of open-loop Stackelberg equilibria in linear-quadratic differential games," *J. Optimization Theory and Applications*, vol. 110, no. 3, pp. 515–544, 2001.

[92] J. Yong, "A leader-follower stochastic linear quadratic differential game," *SIAM Journal on Control and Optimization*, vol. 41, no. 4, pp. 1015–1041, 2002.

[93] A. Bensoussan, S. Chen, and S. P. Sethi, "The maximum principle for global solutions of stochastic Stackelberg differential games," *SIAM*, 2015, to appear.

[94] T. Başar and R. Srikant, "Revenue-maximizing pricing and capacity expansion in a many-users regime," in *Proceedings IEEE INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 1, 2002, pp. 294–301 vol.1.

[95] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Başar, "Dependable demand response management in the smart grid: A Stackelberg game approach," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 120–132, March 2013.

[96] J. Moon and T. Başar, "Estimation over lossy networks: A dynamic game approach," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Florence, Italy, 2013, pp. 2412–2417.

[97] J. Moon and T. Başar, "Minimax estimation with intermittent observations," *Automatica*, Aug. 2015, accepted.

[98] J. Moon and T. Başar, "Control over TCP-like lossy networks: A dynamic game approach," in *Proceedings of American Control Conference*, Washington, DC, June 2013, pp. 1581–1586.

[99] J. Moon and T. Başar, "Control over lossy networks: A dynamic game approach," in *Proceedings of American Control Conference*, Portland, OR, June 2014, pp. 5367–5372.

[100] J. Moon and T. Başar, "Minimax control of MIMO systems over multiple TCP-like lossy networks," in *Proceedings of 19th IFAC World Congress*, Cape Town, South Africa, Aug. 2014, pp. 110–115.

[101] J. Moon and T. Başar, "Minimax control over unreliable communication channels," *Automatica*, vol. 59, pp. 182–193, Sep. 2015.

[102] J. Moon and T. Başar, "Linear-quadratic risk-sensitive mean field games," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, LA, USA, Dec. 2014, pp. 2691–2696.

[103] J. Moon and T. Başar, "Linear quadratic risk-sensitive and robust mean field games," *IEEE Transactions on Automatic Control*, 2015, submitted (under review).

[104] J. Moon and T. Başar, "Linear-quadratic stochastic differential Stackelberg games with a high population of followers," in *Proceedings of the 54th IEEE Conference on Decision and Control*, Osaka, Japan, Dec. 2015, accepted.

[105] J. Moon and T. Başar, "Linear quadratic mean field Stackelberg differential games," *SIAM Journal on Control and Optimization*, 2015, submitted (under review).

[106] T. Başar and P. Bernhard, *H^∞ Optimal Control and Related Minimax Design Problems*, 2nd ed. Boston, MA: Birkhäuser, 1995.

[107] G. Didinsky, "Design of minimax controllers for nonlinear systems using cost-to-come methods," Ph.D. dissertation, University of Illinois, Urbana-Champaign, IL, 1994.

[108] R. Durret, *Probability: Theory and Examples*, 3rd ed. Thomson, 2005.

[109] J. E. Marsden and M. J. Hoffman, *Elementary Classical Analysis*, 2nd ed. Freeman, 1993.

[110] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, 1995.

[111] K. Zhou, *Robust and Optimal Control*. Prentice Hall, 1996.

[112] L. Schenato, "To zero or to hold control inputs with lossy links?" *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1093–1099, May 2009.

[113] J. Huang and M. Huang, "Mean field LQG games with model uncertainty," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 3103–3108.

[114] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. Springer, 2000.

[115] S. Boyd, E. Ghaoui, E. Feron, and B. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*. Philadelphia, PA: SIAM, 1994.

[116] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley & Sons, 1969.

[117] G. Dullerud and F. Paganini, *A Course in Robust Control Theory (A Convex Approach).* Springer, 2000.

[118] P. Nie, L. Chen, and M. Fukushima, "Dynamic programming approach to discrete time dynamic feedback Stackelberg games with independent and dependent followers," *European Journal of Operational Research*, vol. 169, no. 1, pp. 310 – 328, 2006.

[119] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Springer, 1999.

[120] J.-M. Bismut, "Linear quadratic optimal stochastic control with random coefficients," *SIAM Journal on Control and Optimization*, vol. 14, no. 3, pp. 419–444, 1976.

[121] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications.* Springer-Verlag, 1999.

[122] J. L. Doob, *Stochastic Processes.* John Wiley & Sons, 1990.

[123] G. Freiling, "A survey of nonsymmetric Riccati equations," *Linear Algebra and its Applications*, vol. 351-352, pp. 243–270, 2002, Fourth Special Issue on Linear Systems and Control.

[124] U. Vaidya and E. Nicola, "Stabilization of nonlinear systems over packet-drop links: Scalar case," *Systems and Control Letters*, vol. 61, pp. 959–966, 2012.

[125] B.-C. Wang and J.-F. Zhang, "Hierarchical mean field games for multiagent systems with tracking-type costs: Distributed ε-Stackelberg equilibria," *IEEE Transactions on Automatic Control*, vol. 59, no. 8, pp. 2241–2247, Aug 2014.

[126] D. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1.

[127] M. Huang, R. P. Malhamé, and P. E. Caines, "Nash equilibria for large-population linear stochastic systems of weakly coupled agents," in *Analysis, Control and Optimization of Complex Dynamic Systems*, E. K. Boukas and R. P. Malhamé, Eds. New York: Springer, 2005, pp. 215–252.

[128] S. Boucheron, O. Bousquet, and G. Lugosi, "Concentration inequalities," *Advanced Lectures in Machine Learning (Springer)*, pp. 208–240, 2004.