H_∞-Constrained Dynamic Games for Linear Stochastic Systems with Multiple Decision Makers

（複数の意思決定者を伴う線形確率システムにおける H_∞ 制約付き動的ゲーム）

AHMED MOSTAK
D151189

Graduate School of Engineering

Hiroshima University

In Partial Fulfillment of the Requirements For the Degree of

Doctor of Philosophy

September 2018

This thesis is dedicated

to

My mother Monwara Begum and my wife Fatima Tuj Jahra

Acknowledgements

I would like to acknowledge the Japanese Government for its financial support through the MEXT scholarship. It is my pleasure to acknowledge the role of several individuals who helped me complete my Ph.D. research.

I would like to thank my supervisor Professor Dr. Hiroaki Mukaidani for introducing me to the world of control theory research. Professor Mukaidani gave me not only valuable time, but also good advice and a mentor that I will always be grateful for. He helped me hone my skills as a researcher and through his guidance I learned how to complete my research before my deadlines in the most effective way possible.

I would also like to thank to my co-supervisor Dr. Tadashi Shima for helping me understand many advanced mathematical and statistical issues. In addition, some of the basic questions at weekly seminars encourage me to go to the depths of my research.

I would like to thank Professor Hua Xu of the University of Tsukuba for giving me his valuable time and collaborating with me as part of my doctoral degree requirements. He also gave me valuable guidance in writing my Ph.D. dissertation.

Thanks to my sub-advisers, Professor Takio Kurita and Professor Toru Yamamoto for monitoring my research progress regularly and encouraging me. In particular, Professor Takio Kurita arranged monthly joint seminars, which were very fruitful for me.

I am very grateful to the reviewer Professor Nobutaka Wada for reviewing my thesis appropriately from beginning to end. Professor Wada has put forward a number of suggestions for improving my thesis. In addition, some of his noble ideas for my future research will make me remember him.

I would like to thank my research colleagues for maintaining a good research environment in the laboratory. In particular, Mr. Kawakami took part in the weekly seminars and made them more vivid, interesting and meaningful for me.

Abstract

In this thesis, an H_∞-constrained incentive Stackelberg games for stochastic systems with deterministic external disturbances are investigated. As part of the preliminary study, some results of the deterministic system are also presented. Although we focus on continuous-time, we have studied dynamic games for both discrete- and continuous-time systems. In the case of discrete-time, both deterministic and stochastic systems are investigated. One leader and multiple followers are considered for both finite- and infinite-time cases. We also studied multiple leaders and multiple followers for a continuous-time stochastic system. To simplify the calculation, only the infinite-time horizon of continuous-time is emphasized.

In the incentive Stackelberg game, players are divided into two categories; the leader group and the follower group. For a single leader game, incentive Stackelberg strategy is an extensive idea in which the leader can achieve his team-optimal solution in a Stackelberg game. Multiple leaders and multiple followers have made the game more complex and challenging. In the leaders' group and followers' group, players are supposed to be non-cooperativ; subsequently, Nash equilibrium is investigated. Several theorems and lemmas are designed to study the incentive Stackelberg game problems. For multiple leader games, an incentive structure is developed in such a way that leaders achieve Nash equilibrium by attenuating the disturbance under H_∞-constrained. Simultaneously, followers achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance is considered. The deterministic disturbances and their attenuation to stochastic systems under the H_∞-constrained is one of the main attractions of this thesis. Problems involving deterministic disturbance must be attenuated at a given target called disturbance attenuation level $\gamma > 0$. Surprisingly, the concept of solving the disturbance reduction problem under the H_∞-constrained seems like a Nash equilibrium between the disturbance input and the control input.

In this research, a very general and simple linear stochastic system governed by Itô differential equation has been studied. This thesis studies the most common linear quadratic (LQ) optimal control in the game problems. In order to solve the LQ problem, stochastic dynamic programming (SDP) and stochastic maximum principle [Peng (1990)] are used. Cooperative game problems and non-cooperative game problems are solved based on the concepts of Pareto optimality and Nash equilibrium solutions, respectively. Several basic problems are completely solved and useful for current research. The main task to solve the LQ problem is to find a matrix solution of algebraic Riccati equations (AREs). Newton's method and Lyapunov iterative method are used to solve such AREs.

However, the main objective of this research is to investigate the incentives Stackelberg strategy, preliminary research and synthesis of LPV systems for multiple decision makers. We aim to better understand to implement our current idea for LPV system in the future. H_∞-constrained Pareto optimal strategy for stochastic linear parameter varying (LPV) systems with multiple decision makers is investigated. The modified stochastic bounded real lemma and linear quadratic control (LQC) for the stochastic LPV systems are reformulated by means of linear matrix inequalities (LMIs). In order to decide the strategy-set of multiple decision makers, Pareto optimal strategy is considered for each player and the H_∞-constrained is imposed. The solvability conditions of the problem are established from cross-coupled matrix inequalities (CCMIs). Several academic and real-life numerical examples have also been resolved to demonstrate the usefulness of our proposed schemes.

This thesis consists of seven chapters. In Chapter 1, the research background, motivation, research survey, objectives and outlines of the thesis are described. Some basic definitions and preliminary results are also introduced in this chapter. Chapter 2 of the thesis summarizes some of the preliminary mathematical problems based on discrete-time and continuous-time stochastic optimal control. The exogenous disturbance problem and its attenuation of the H_∞-constrained are presented. In Chapter 3, the incentive Stackelberg game for a discrete-time deterministic system is considered. It explains two levels of hierarchy with one leader and multiple followers. Followers are supposed to act non-cooperatively. Exogenous disturbance also exists in the system and is attenuated under the H_∞-constrained. Chapter 4 investigates the incentive Stackelberg game for discrete-time stochastic systems. The structure of the game is very similar to Chapter 3. It is a single leader and multiple non-cooperative followers with exogenous disturbance which is attenuated under the H_∞-constrained in the 2-level hierarchy. Therefore, Chapter 4 can be viewed as the stochastic version of the deterministic game described in Chapter 3. In Chapter 5, the continuous-time incentive Stackelberg games for multiple leader and multiple followers are investigated. The external disturbance is included with the system, as usual. The information pattern of the game is more complex than before. Each leader must achieve Nash equilibrium and use the H_∞-constrained to reduce the external disturbance. Each leader separately announces the declares incentive Stackelberg strategies for each follower. Each follower employs a leader incentive mechanism that follows the Nash equilibrium in a follower group. Leaders and followers do not cooperate with their group. Chapter 6 discusses the Pareto optimal strategy for stochastic LPV system with multiple decision makers. In the dynamic game of uncertain stochastic systems, multiple participants can be used for more realistic plants. The system includes disturbances that are attenuated under the H_∞-constrained. Finally, in Chapter 7, the thesis is concluded with some motivating guidelines for future research.

In this thesis, two appendices are included. The Appendix A discusses how to solve convex optimization problems using linear matrix inequalities (LMIs) and special cases to solve systems and control theory problems. Some preliminary results on static output feedback optimal control are given in Appendix B. Here we consider the linear quadratic optimal cost control problem for stochastic Itô differential equations. Several definitions, theorems, and lemmas are studied for future research. To solve the output feedback control problem, Newton's algorithm and corresponding MATLAB codes are already developed. Numerical examples of a very basic problem have been solved. The problem is already formulated for future investigation.

Contents

List of Figures

Chapter 1

Introduction

Game theory is a mathematical model for studies of conflict and cooperation among intelligent rational decision makers. Game theory is mainly applied to economics, politics, psychology, logic, and computer science. Initially, it involved a zero-sum game in which one's earnings led to the loss of other. Nowadays game theory applies to a wide range of behavioral relationships. It is a general term for the logical decision science of humans, animals, and computers. Game theory is related to strategic interactions among multiple decision makers and named *players*. Every player has a nontrivial strategy chosen based on the payoff, i.e., each player has an objective function to maximize, called *profit* function, or to minimize, called *cost* function. In a multi-player game, each player's objective function consists of at least one other player's choices called *decision variables*. The decision variable is the amount managed by the decision maker. A decision variable that determines the value that generates the optimum value of the objective function. The *cooperative game* investigates the relative amount of power each player holds in these alliance games, or how the alliance allocates their returns. This applies to what happens in political science or international relations and the concept of power. On the other hand, if cooperation is not allowed between players, we call the game *non-cooperative game*. The non-cooperative game concerns the analysis of strategic choices. In this case, two or more players cannot move together from the solution point. The solution point where players can take benefit from unilateral movements is called *Nash Equilibrium*, and is named after John Nash. A non-cooperative game is called *zero-sum* if the sum of the players' objective functions equals zero. The zero-sum game is a mathematical representation in which each player's utility gain or loss is completely balanced with the loss or gain of other player's utility. If the players' total returns add up and subtract the total loss, they will total zero. Similarly, we can define *nonzero-sum* and *constant sum* game. If the player's behavior uniquely determines the result captured by the objective function, the game is considered *deterministic*. In deterministic games, player action solutions produce completely predictable results. On the other hand, if at least one player's objective function is a known probability distribution of an additional variable (state), then, it becomes *stochastic game*. The stochastic game is a dynamic game with a stochastic transition to be played by one or more players. If no player can get any information about the behavior of any other player, it is called a *dynamic game*. On the other hand, if players can only access the information shared by all, we say the game is *static*. If the evolution of the decision-making process (controlled by the player over time) occurs in a continuous period of time and usually involves a differential equation, the dynamic game is considered a *differential game*. If it occurs on a discrete-time horizon case, a dynamic game is called a *discrete time game*.

1.1 Research background and motivation

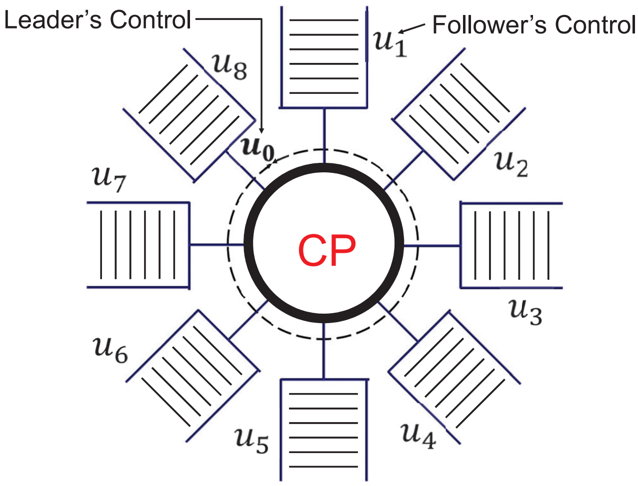


Fig. 1.1: Packet switch in a loop architecture.

The engineering application of the incentive Stackelberg strategy is a scheduling problem of packet switches working in the loop structure (Figure 1.1) introduced by [Saksena and Cruz (1985)]. Communication in high-speed networks can be switched optically or electronically. Although optical switches are advantageous for circuit switching, it is generally considered difficult to combine them with packet switching. In packet switching and scheduling, the switch provides lossless communications for sessions with certain smoothing attributes and allows the use of input flow control to translate the session into a smooth session. When switching between connections according to a scheduling strategy, inbound packets on each connection are stored in a limited-capacity buffer and managed by the central processor. Packets are rejected, when the buffer is full. When the central processor serves a particular connection, the buffering force of the connection is controlled locally based on the state information of the buffer. The goal of the local controller is to maximize the data transmission on the connection. On the other hand, the central processor knows the state of all connection buffers and all local controllers' control actions. The goal of the central processor is to use scheduling policies to maximize the total data transmission from all connections. In this problem, the central processor is represented as a leader and the local connection controller is represented as a follower. The problem information structure allows the leader to access the follower's decision value and observations at each stage of the process. However, our motivation is that the incentive Stackelberg strategy for the above problem comes from static games, our goal is to deal with more challenging dynamic games. Furthermore, dynamic games using the stochastic system make the thesis more informative.

1.2 Stochastic differential equations

Since the problem investigated in this thesis refers to a stochastic control system, we recall some facts on the stochastic differential equations.

Definition 1.1. *[Dragan et al. (2006)] A measurable space is an ordered pair $(X, \mathscr{A})$, in which X is a set and $\mathscr{A}$ is a σ-algebra of subsets of X, that is, $\mathscr{A}$ is a family of subsets $A \subset X$ such that*

(i) $X \in \mathscr{A}$;

(ii) *If* $A \in \mathscr{A}$, *then* $X - A \in \mathscr{A}$;

(iii) *If* $A_n \in \mathscr{A},\ n \geq 1$, *then* $\cup_{n=1}^{\infty} A_n \in \mathscr{A}$.

Definition 1.2. *[Dragan et al. (2006)] For a measurable space $(X, \mathscr{A})$, a function $\mu : \mathscr{A} \to [0, \infty]$ is called a measure if:*

(i) $\mu(\emptyset) = 0$

(ii) *if* $A_n \in \mathscr{A},\ n \geq 1$ *and* $A_i \cap A_j = \emptyset$ *for* $i \neq j$, *then*

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

A triplet $(X, \mathscr{A}, \mu)$ is called a space with measure. If $\mu(X) = 1$, then $\mu = P$ is a probability on $\mathscr{A}$ and the triplet $(X, \mathscr{A}, P)$ is called a probability space.

Definition 1.3. *[Dragan et al. (2006)] An r-dimensional stochastic process* $x(t),\ t \in [0,\infty)$ *is called a stochastic process with independent increments if for all* $0 = t_0 < t_1 < \cdots < t_N = t$*, the random vectors* $x(t_0),\ x(t_1) - x(t_0),\ \ldots\ ,x(t_N) - x(t_{N-1})$ *are independent.*

Definition 1.4. *[Dragan et al. (2006)] A stochastic process* $w = \{w(t)\}_{t\in[0,\infty)}$ *in a probability space* $(X,\ \mathscr{A},\ P)$ *is called a standard Brownian motion or a standard Wiener process if:*

(i) $w(0) = 0$*;*

(ii) *with probability* 1*, the function* $t \to w(t)$ *is continuous in* t*;*

(iii) $w(t)$ *is a stochastic process with independent increments and the increment* $w(t+s) - w(s)$ *has the normal distribution* $N(0,\ t)$ *with* $t,\ s \in [0,\ \infty)$*;*

(iv) $\mathbb{E}[w(t)] = 0,\ t \in [0,\ \infty),\ \mathbb{E}[|w(t) - w(s)|^2] = |t - s|$ *with* $t,\ s \in [0,\ \infty)$.

1.2.1 Stochastic integrals

Through the one dimensional Wiener process $w(t) \in \mathbb{R}$, let us define a continuous sample path as a solution to the following stochastic differential equation:

$$dx(t) = f(t,\ x(t))dt + g(t,\ x(t))dw(t), \quad x(0) = x_0, \tag{1.1}$$

for all $t \in [0,\ \infty),\ x \in \mathbb{R}^n$. Basically, the stochastic differential equation (1.1) is the representation of the following integral sign.

$$x(t) = x_0 + \int_0^t f(s,\ x(s))ds + \int_0^t g(s,\ x(s))dw(s). \tag{1.2}$$

The last integral term on the right side of equation (1.2) is called a stochastic integral.

For a suitable function g [Higham (2001)], the integral $\int_0^T g(t)dt$ can be approximated by Riemann-Stieltjes integral with the sum

$$\sum_{k=0}^{N-1} g(t_k)(t_{k+1} - t_k), \tag{1.3}$$

where the discrete points $t_k = k\delta t$ were already introduced. The Riemann sum in (1.3) is based on left end-point. In fact, the integral is defined by taking the limit $\delta t \to 0$ in (1.3). Using a similar idea, we can consider the sum of the form

$$\sum_{k=0}^{N-1} g(t_k)(w(t_{k+1}) - w(t_k)). \tag{1.4}$$

By analogy with (1.3), it can be considered as an approximation of the stochastic integral $\int_0^T g(t)dw(t)$. This is known as the Itô integral. Here, we integrate g with respect to the Wiener process $w(t)$.

Remark 1.1. *The trajectory of the Wiener process has an infinite change in any interval* $[0,\ T]$, $T > 0$. *In other words, for any* $\delta > 0$,

$$P\left[\sum_{k=0}^{N-1} |w(t_{k+1}) - w(t_k)| > \delta\right] \to 1, \tag{1.5}$$

holds as $\max(t_{k+l} - t_k) \to 0,\ 0 = t_0 < t_l < \ldots < t_N = T$. *It should be noted that because of* (1.5)*, the Itô stochastic integral can fail to exist in the sense of Riemann-Stieltjes [Afanasiev et al. (2013)].*

An alternative approach to the Riemann sum (1.3) based on mid-point can be written as

$$\sum_{k=0}^{N-1} g\left(\frac{t_k + t_{k+1}}{2}\right)(t_{k+1} - t_k). \tag{1.6}$$

The corresponding alternative to (1.4) is

$$\sum_{k=0}^{N-1} g\left(\frac{t_k + t_{k+1}}{2}\right)(w(t_{k+1}) - w(t_k)). \tag{1.7}$$

The 'mid-point' sum (1.7) is known as Stratonovich integral.

1.2.2 Itô's formula

Let us define a scalar function $V(t,\ x)$ for which the partial derivatives V_t, V_x and V_{xx} exist. If the process $x(t)$ possesses the following stochastic differential equation:

$$dx(t) = f(t)dt + \sigma(t)dw(t). \tag{1.8}$$

then the process $\theta(t) = V(t, x(t))$ also has a stochastic differential $d\theta(t)$, defined by the following formula:

$$\begin{aligned} d\theta(t) = \Big[&V_t(t,\ x(t)) + V_x^T(t,\ x(t))f(t) \\ &+ \frac{1}{2}\mathbf{Tr}[\sigma(t)\sigma^T(t)V_{xx}(t,\ x(t))]\Big]dt + V_x^T(t,\ x(t))\sigma(t)dw(t), \end{aligned} \tag{1.9}$$

where **Tr** is the trace of a matrix; $V_x \in \mathbb{R}^n$ is the vector with components $\frac{\partial V}{\partial x_i}$; V_{xx} is the square matrix with elements $\frac{\partial^2 V}{\partial x_i \partial x_j}$, $(i,\ j = 1,\ \ldots\ ,N)$ [Afanasiev et al. (2013)].

Proof. Let us consider two arbitrary moments τ_1 and τ_2 $(\tau_1 < \tau_2)$. Let us divide the interval $[\tau_1,\ \tau_2]$ into some sub-intervals as follows:

$$\tau_1 = t_0 < t_1 < \cdots < t_N = \tau_2.$$

Therefore,

$$\theta(\tau_2) - \theta(\tau_1) = \sum_{i=0}^{N-1} [V(t_{i+1},\, x(t_{i+1})) - V(t_i,\, x(t_i))]. \tag{1.10}$$

Applying Taylor's formula on the term of right hand side, we obtain

$$\begin{aligned} V(t_{i+1},\, x(t_{i+1})) - V(t_i,\, x(t_i)) & \\ = & V_t(t_i + \alpha_i(t_{i+1} - t_i),\, x(t_i))(t_{i+1} - t_i) \\ & + V_x^T(t_i,\, x(t_i))(x(t_{i+1}) - x(t_i)) \\ & + \frac{1}{2}\mathbf{Tr}\Big[V_{xx}(t_i,\, x(t_i) + \lambda_i(x(t_{i+1}) - x(t_i))) \\ & \times (x(t_{i+1}) - x(t_i))(x(t_{i+1}) - x(t_i))\Big], \end{aligned}$$

where, α_i and λ_i are some numbers form the interval $[0,\ 1]$. The increment of the process $x(t)$ of the equation (1.8) can be represented by

$$x(t_{i+1}) - x(t_i) = f(t_i)(t_{i+1} - t_i) + \sigma(t_j)(w(t_{i+1}) - w(t_i)), \tag{1.11}$$

with $\max[t_{i+1} - t_i] \to 0$. Therefore,

$$\begin{aligned} V(t_{i+1},\, x(t_{i+1})) - V(t_i,\, x(t_i)) & \\ = & [V_t(t_i,\, x(t_i)) + V_x^T(t_i,\, x(t_i)) f(t_j)](t_{i+1} - t_i) \\ & + V_x^T(t_i,\, x(t_i))\sigma(t_j)](w(t_{i+1}) - w(t_i)) \\ & + \frac{1}{2}\mathbf{Tr}[V_{xx}(t_i,\, x(t_i))\sigma(t_i)\sigma^T(t_i)(t_{i+1} - t_i)]. \end{aligned}$$

Finally, from (1.10) we can obtain,

$$\begin{aligned} \theta(t_2) - \theta(t_1) = \sum_{i=0}^{N-1} \Bigg[\Big[& V_t(t_i,\, x(t_i)) + V_x^T(t_i,\, x(t_i)) f(t_i) \\ & + \frac{1}{2}\mathbf{Tr}[V_{xx}(t_i,\, x(t_i))\sigma(t_i)\sigma^T(t_i)]\Big](t_{i+1} - t_i) \\ & + V_x^T(t_i,\, x(t_i))\sigma(t_j)](w(t_{i+1}) - w(t_i))\Bigg]. \end{aligned}$$

By observing the limits of all sums in this expression, it becomes as follows:

$$\begin{aligned} d\theta(t) = \Big[& V_t(t,\, x(t)) + V_x^T(t,\, x(t)) f(t) \\ & + \frac{1}{2}\mathbf{Tr}[\sigma(t)\sigma^T(t)V_{xx}(t,\, x(t))]\Big] dt + V_x^T(t,\, x(t))\sigma(t) dw(t), \end{aligned} \tag{1.12}$$

Hence, Itô's formula is proved. □

1.2.3 Stochastic differential games

Stochastic differential games are a type of decision-making problem in which the evolution of states is described by stochastic differential equations and the players work throughout a time interval [Başar and Olsder (1999)]. An *N-person stochastic differential game* within a fixed duration involves the following:

(i) An index set $\mathcal{N} = \{1, \ldots, N\}$ is called *players' set*.

(ii) The prior information about time, $[0, T]$ is called time domain.

(iii) An infinite set $S_0 \in \mathbb{R}^n$ with some topological structure containing elements $\{x(t) \in S_0, 0 \le t \le T\}$ is called the game's *trajectory space*. Its elements $\{x(t), 0 \le t \le T\}$ constitutes the game's allowable state trajectory. Furthermore, for each fixed $t \in [0, T]$, $x(t) \in S_0$.

(iv) An infinite set $S_i \in \mathbb{R}^{m_i}$ with some topological structure, defined for each $i \in \mathcal{N}$ called the *control (action) space* of i-th player, P_i, whose elements $\{u_i(t), 0 \le t \le T\}$ are the controls of P_i. Furthermore, for each fixed $t \in [0, T]$, $u_i(t) \in S_i$.

(v) A stochastic differential equation

$$dx(t) = f(t, x(t), u_1(t), \ldots, u_N(t))dt + \sigma(t, x(t))dw(t), \quad x(0) = x_0, \tag{1.13}$$

whose solution represents the *state trajectory* of the game. Here, $\sigma(t, x(t))$ is an $m \times \theta$ matrix and $w(t)$ is a θ dimensional Wiener process.

(vi) A set of value functions $\eta_i(\cdot)$ defined for each $i \in \mathcal{N}$

$$\eta_i(t) = \{x(s), 0 \le s \le \varepsilon_t^i\}, \quad 0 \le \varepsilon_t^i \le t, \tag{1.14}$$

where ε_t^i is non-decreasing in t, and $\eta_i(t)$ determines the state information gained and recalled by P_i at time $t \in [0, T]$. Specification of $\eta_i(\cdot)$ characterizes the *information structure/pattern* of P_i.

(vii) A sigma-field N_t^i, in S_0, generated for each $i \in \mathcal{N}$ by the set $\{x(s) \in B \mid 0 \le s \le \varepsilon_t^i\}$, where $B \subset S_0$ is a Borel set - is called the *information field* of P_i.

(viii) A pre-specified class Γ_i of mappings

$$\gamma_i : [0, T] \times S_0 \to S_i,$$

with the property that $u_i(t) = \gamma_i(t, x)$ is N_t^i-measurable (i.e. it is adapted to the information field N_t^i). Γ_i is the strategy space of P_i and each of its elements γ_i is a permissible strategy for P_i.

(ix) Two functions $q_i : S_0 \to \mathbb{R},\ g_i : [0,\ T] \times S_0 \times S_1 \times \cdots \times S_N \to \mathbb{R}$ defined for each $i \in \mathcal{N}$, so that the composite functional

$$J_i(u_1, \ldots,\ u_N) = \mathbb{E}\left[\int_0^T g_i(t,\ x(t),\ u_1(t), \ldots,\ u_N(t))dt + q_i(x(T))\right]. \tag{1.15}$$

is well defined for every $u_j(t) = \gamma_j(t,\ x),\ \gamma_j \in \Gamma_j (j \in \mathcal{N})$, and for each $i \in \mathcal{N}$, J_i is the *cost functional* of P_i in a fixed duration stochastic differential game and $\mathbb{E}[\cdot]$ is the expectation operator.

1.2.4 Optimal control

The optimal control is a specific branch of modern control to provide an attractive analysis design. The final result of an optimal design is not considered to be stable, having a certain bandwidth, or satisfying any kind of ideal constraints related to classical control. However, it is considered to be the best type of a particular system – therefore, the term optimal. Linear optimal control is a special kind of optimal control in which the controlled device is assumed to be linear, and the controller, which generates the optimal control, is limited to linear. A linear controller operating with a quadratic performance optimization index is called a linear quadratic (LQ) method. We focus on linear quadratic control problems where the cost functional is quadratic and the state equation is linear. The control theory of deterministic systems strongly influences the research of stochastic optimal control problems, in which the state of the system is represented by a stochastic process. In the long history of stochastic systems studies, the class associated with white noise perturbations attracted a lot of attention to the control literature. The goal of optimization is very common because it can be viewed in different ways depending on the method.

It is worth mentioning that this thesis only studies the convex optimization problem. That is, the weighting matrix of all linear quadratic costs in this thesis is assumed to be positive definite/positive semi-definite.

Consider the following stochastic LQ system

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{1.16a}$$

$$J(x^0;\ u) = \mathbb{E}\left[\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)\ dx\right], \tag{1.16b}$$

with $Q \geq 0$ and $R > 0$.

According to [Rami and Zhou (2000)], let us define a subset $\mathcal{I}$ of $\mathcal{S}^n$

$$\mathcal{I} \triangleq \{X \in \mathcal{S}^n \mid \det[R] \neq 0\}. \tag{1.17}$$

It should be noted that $\mathcal{I} \neq \phi$ when R is nonsingular.

Let is define the following stochastic algebraic Riccati operator $\mathcal{R} : \mathcal{I} \to \mathcal{S}^n$ by

$$\mathcal{R}(X) \triangleq A^T X + XA + A_p^T X A_p + Q - XBR^{-1}B^T X. \tag{1.18}$$

Moreover, let us introduce the following subset $\mathcal{P}$ of $\mathcal{S}^n$ as

$$\mathcal{P} \triangleq \{P \in \mathcal{S}^n \mid \mathcal{R}(P) \geq 0,\ R > 0\}. \tag{1.19}$$

By applying Schur's lemma, we can write (1.19) as the following LMI format:

$$\mathcal{P} \triangleq \{P \in \mathcal{S}^n \mid \mathcal{M}(P) \geq 0,\ R > 0\}, \tag{1.20}$$

where

$$\mathcal{M}(X) \triangleq \left[\begin{array}{c|c} A^T X + XA + A_p^T X A_p + Q & XB \\ \hline B^T X & R \end{array}\right]. \tag{1.21}$$

$\mathcal{P}$ is then seen to be convex as $\mathcal{M}$ is affine [Rami and Zhou (2000)].

1.2.5 Non-cooperative games

In order to express non-cooperative games, the following objects are necessary:

- the number of players,
- the actions that each player may take and the restrictions imposed on them,
- objective function for each player to optimize,
- if the player is allowed to perform multiple actions, the time sequence of the actions,
- the information pattern and at what point each player can use information based on past actions of other players,
- whether the player's behavior is the result of a fixed (known) distribution of stochastic events or not.

Accordingly, we consider an N-player game, with $P_1, \ldots, P_N$ denoting the Players set. The decision or action variable of Player i is denoted by $x_i \in X_i$, where X_i is the action set of Player i. The action set could be a finite set, for example, with $N = 2$, we could have a coupled constraint set described by: $0 \leq x_1, x_2 \leq 1,\ x_1 + x_2 \leq 1$. If we consider the objective/cost function of Player i will be denoted by $J_i(x_i, x_{-i})$, where x_{-i} stands for the action variables of all players except the i-th one.

Let $\Omega \in X$ be the constraint set where the actions variables are feasible. Now, an N-tuple of action variables $x^* \in \Omega$ constitutes a *Nash equilibrium* (or, non-cooperative equilibrium) (NE) if, for all $i \in N$,

$$J_i(x_i^*, x_{-i}^*) \leq J_i(x_i, x_{-i}^*), \qquad \forall\, x_i \in X_i, \text{ such that } (x_i, x_{-i}^*) \in \Omega. \tag{1.22}$$

If $N = 2$, and $J_1 = -J_2 = J$, then we have a two-player zero-sum game (ZSG), with Player 1 minimizing J and Player 2 maximizing the same quantity. In this case, the Nash equilibrium becomes the *saddle-point equilibrium*(SPE)

$$J(x_1^*, x_2) \leq J(x_1^*, x_2^*) \leq J(x_1, x_2^*), \qquad \forall (x_1, x_2) \in X. \tag{1.23}$$

This also implies that the order in which minimization and maximization are carried out is inconsequential, that is

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} J(x_1, x_2) = \max_{x_2 \in X_2} \min_{x_1 \in X_1} J(x_1, x_2) = J(x_1^*, x_2^*) =: J^*. \tag{1.24}$$

We then say in this case that *the zero-sum game does not have a saddle point in pure strategies* if

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} J(x_1, x_2) > \max_{x_2 \in X_2} \min_{x_1 \in X_1} J(x_1, x_2). \tag{1.25}$$

This opens the door for looking for a *mixed-strategy equilibrium*.

With just replacing x_i's by p_i's, where $p_i \in \mathscr{P}_i$ is the set of all probability distributions on X_i, a pair (p_1^*, p_2^*) constitutes a *mixed-strategy saddle-point equilibrium* (MSSPE), if

$$J(p_1^*, p_2) \leq J(p_1^*, p_2^*) \leq J(p_1, p_2^*), \qquad \forall (p_1, p_2) \in \mathscr{P}_1 \times \mathscr{P}_2, \tag{1.26}$$

where

$$J(p_1, p_2) = \mathbb{E}_{p_1, p_2}[J(x_1, x_2)].$$

Similarly, if there exists no Nash equilibrium for an N-player game, the n-tuple $(p_1^*, \ldots, p_N^*)$ is in mixed-strategy Nash equilibrium (MSNE) if

$$J_i(p_i^*, p_{-i}^*) \leq J_i(p_i, p_{-i}^*), \qquad \forall p_i \in \mathscr{P}_i. \tag{1.27}$$

A precise definition of extensive form of a dynamic game now follows.

Definition 1.5. *Extensive form of an N-person nonzero-sum finite game without chance moves is a tree structure with* *(i)* *a specific vertex indicating the starting point of the game,*

(ii) *N cost functions, each one assigning a real number to each terminal vertex of the tree, where the i-th cost function determines the loss to be incurred to P_i,*

(iii) *a partition of the nodes of the tree into N player sets,*

(iv) *a subpartition of each player set into information sets $\{\eta_j^i\}$, such that the same number branches emanate from every node belonging to the same information set and no node follows another node in the same information set.*

1.2.6 Cooperative games

In the cooperative game, the strategies of individual players are concentrated in the coalition. If the players are able to reach a cooperation agreement so that the choice of action or decision can be made collectively and with complete trust, the game is called a cooperative. One of the main problems in the cooperative game is how to fairly distribute the big coalition rewards among players. If the solution can be found in vector form, it can represent the task of each player. In this thesis, we propose the Pareto concept, which was named after the economist Vilfredo Pareto, for the cooperative games. Let us consider the following linear stochastic system with linear quadratic cost functions:

$$dx(t) = [Ax(t) + \sum_{i=1}^{N} B_i u_i(t)]dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.28a}$$

$$J_i(x_0, u_1, \ldots, u_N) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left[x^T(t) Q_i x(t) + \sum_{j=1}^{N} u_j^T(t) R_{ij} u_j(t)\right] dt\right], \tag{1.28b}$$

where $Q_i = Q_i^T \geq 0$, $R_{ij} = R_{ij}^T \geq 0$ for $i \neq j$ and $R_{ii} = R_{ii}^T > 0$, $i, j = 1, \ldots, N$.

Definition 1.6. *A strategy-set $(u_1, \cdots, u_N)$ is said to be a Pareto optimal strategy if it minimizes a sum of the cost of functional of all players denoted by*

$$J(u_1, \cdots, u_N) = \sum_{i=1}^{N} r_i J_i(x_0, u_1, \cdots, u_N), \tag{1.29}$$

where $\sum_{i=1}^{N} r_i = 1$ for some $0 < r_i < 1$.

Theorem 1.1. *For the stochastic optimal control problem* (1.28)*, the optimal linear feedback strategy for the a Pareto game is given by*

$$u^*(t) = -\mathbf{R}^{-1}\mathbf{B}^T P x(t). \tag{1.30}$$

where $P^T = P \geq 0$ is the solution of the following stochastic algebraic Riccati equation (SARE):

$$PA + A^T P + \mathbf{Q} - P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P + A_p^T P A_p = 0 \tag{1.31}$$

with

$$\begin{aligned}
\mathbf{B} &:= \left[B_1, \ \ldots \ ,B_N\right], \\
u(t) &:= \begin{bmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{bmatrix}, \\
\mathbf{Q} &:= \sum_{i=1}^{N} r_i Q_i, \\
\mathbf{R} &:= \textbf{block diag}\left[\sum_{i=1}^{N} r_i R_{i1} \quad \ldots \quad \sum_{i=1}^{N} r_i R_{iN}\right].
\end{aligned}$$

Proof. If we centralized the system (1.28) base on the Definition 1.6, we can can rewrite it as follows:

$$dx(t) = [Ax(t) + \mathbf{B}u(t)]dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.32a}$$

$$J(x_0, \ u) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)\mathbf{Q}x(t) + u^T(t)\mathbf{R}u(t)\right) dt\right], \tag{1.32b}$$

where

$$\begin{aligned}
\mathbf{B} &:= \left[B_1, \ \ldots \ ,B_N\right], \\
u(t) &:= \begin{bmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{bmatrix}, \\
\mathbf{Q} &:= \sum_{i=1}^{N} r_i Q_i, \\
\mathbf{R} &:= \textbf{block diag}\left[\sum_{i=1}^{N} r_i R_{i1} \quad \ldots \quad \sum_{i=1}^{N} r_i R_{iN}\right].
\end{aligned}$$

With the control (1.30) the system (1.32) becomes

$$dx(t) = [A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]x(t)dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.33a}$$

$$\begin{aligned}
J(x_0, \ u^*) &:= \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)\mathbf{Q}x(t) + x^T(t)P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T Px(t)\right) dt\right], \\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\infty x^T(t)[\mathbf{Q} + P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]x(t)dt\right].
\end{aligned} \tag{1.33b}$$

Let $V(x(t)) = x(t)^T Px(t)$ be the Lyapunov candidate for the system (1.32), where P is a symmetric positive semi-definite matrix. Now applying Itô's formula, we obtain

$$dV(x(t)) = V_x[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]x(t) + \frac{1}{2}x^T(t)A_p V_{xx} A_p x(t)$$

$$= x^T(t)[[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]^T P + P[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P] + A_p^T P A_p]x(t), \tag{1.34}$$

which is stable if $dV(x(t)) < 0$. If $(A,\ A_p \mid C)$ is exactly observable, we can form the Lyapunov stabilizable equation. By integrating and taking expectoration operator ($\mathbb{E}[\cdot]$) in (1.34) as follows:

$$\mathbb{E}\left[\int_0^\infty x^T(t)[[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]^T P + P[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P] + A_p^T P A_p]x(t)dt\right]$$
$$= -\mathbb{E}\left[\int_0^\infty x^T(t)[\mathbf{Q} + P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]x(t)dt\right], \tag{1.35}$$

i.e.,

$$[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P]^T P + P[A - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P] + A_p^T P A_p + \mathbf{Q} + P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P = 0. \tag{1.36}$$

After simplification we can find

$$PA + A^T P + \mathbf{Q} - P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P + A_p^T P A_p = 0. \tag{1.37}$$

Hence, the first part of Theorem 1.1 is proved. □

1.2.7 Nash games

Definition 1.7. *An N-tuple of strategies*

$$u_* := \{u_1^*, u_2^*,\ \dots,\ u_N^*\},$$

with $u_i^* \in U_i,\ i \in \mathbb{N}$ constitutes a Nash equilibria for an N-person nonzero-sum finite game, if the following N inequalities are satisfied for all $u_i \in U_i,\ i \in \mathbb{N}$:

$$J_i^* := J_i(x^0,\ u_1^*,\ \dots,u_N^*) \le J_i(x^0,\ u_1^*,\ \dots,\ u_{i-1}^*,\ u_i,\ u_{i+1}^*,\ \dots,u_N^*). \tag{1.38}$$

where The N-tuple of quantities $\{J_1^*,\ \dots\ ,J_N^*\}$ is known as a Nash equilibrium outcome of the nonzerosum finite game.

Let us consider the following linear stochastic system with linear quadratic cost functions,

$$dx(t) = [Ax(t) + \sum_{i=1}^N B_i u_i(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{1.39a}$$

$$J_i(x^0,\ u_1,\ \dots,u_N) = \mathbb{E}\left[\int_0^\infty [x^T(t)Q_i x(t) + u_i^T(t)R_i u_i(t)]dt\right], \tag{1.39b}$$

where with $Q_i = Q_i^T \ge 0$ and $R_i = R_i^T > 0$, $i = 1,\ \dots\ ,N$.

Theorem 1.2. *The optimal linear feedback strategies for the Nash games are given by*

$$u_i^*(t) = K_i x(t) = -R_i^{-1} B_i^T P_i x(t). \tag{1.40}$$

where $P_i^T = P_i \geq 0$ is the solution of the following stochastic algebraic Riccati equation (SARE):

$$P_i \left(A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T \right) + \left(A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T \right)^T P_i \\ + Q_i - P_i B_i R_i^{-1} B_i^T P_i + A_p^T P_i A_p = 0. \tag{1.41}$$

Proof. The stochastic system (1.39a) can be rewritten for i-th player considering others players optimal strategy as follows:

$$dx(t) = \left[\left(A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T \right) x(t) + B_i u_i(t) \right] dt + A_p x(t) dw(t), \quad x(0) = x^0, \\ = [\mathbf{A} x(t) + B_i u_i(t)]\, dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.42}$$

where

$$\mathbf{A} = A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T.$$

Let $V(x(t)) = x(t)^T P_i x(t)$ be the Lyapunov candidate for the system (1.39), where P_i is a symmetric positive semi-definite matrix. Now applying Itô's formula, we obtain

$$dV(x(t)) = V_x[\mathbf{A} - B_i R_i^{-1} B_i^T P] x(t) + \frac{1}{2} x^T(t) A_p V_{xx} A_p x(t) \\ = x^T(t)[[\mathbf{A} - B_i R_i^{-1} B_i^T P_i]^T P_i + P_i[\mathbf{A} - B_i R_i^{-1} B_i^T P_i] + A_p^T P_i A_p] x(t), \tag{1.43}$$

which is stable if $dV(x(t)) < 0$. If $(\mathbf{A},\ A_p \mid C)$ is exactly observable, we can form the Lyapunov stabilizable equation. By integrating and taking expectoration operator ($\mathbb{E}[\cdot]$) in (1.43) as follows:

$$\mathbb{E}\left[\int_0^\infty x^T(t)[[\mathbf{A} - B_i R_i^{-1} B_i^T P_i]^T P_i + P_i[\mathbf{A} - B_i R_i^{-1} B_i^T P_i] + A_p^T P_i A_p] x(t) dt \right] \\ = -\mathbb{E}\left[\int_0^\infty x^T(t)[Q_i + P_i B_i R_i^{-1} B_i^T P_i] x(t) dt \right], \tag{1.44}$$

i.e.,

$$[\mathbf{A} - B_i R_i^{-1} B_i^T P_i]^T P_i + P_i[\mathbf{A} - B_i R_i^{-1} B_i^T P_i] + A_p^T P_i A_p + Q_i + P_i B_i R_i^{-1} B_i^T P_i = 0. \tag{1.45}$$

After simplification we can find

$$P_i\mathbf{A} + \mathbf{A}^T P_i + Q_i - P_i B_i R_i^{-1} B_i^T P_i + A_p^T P_i A_p = 0, \tag{1.46}$$

or,

$$P_i \left(A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T \right) + \left(A - \sum_{j=1,\ j\neq i}^{N} B_j R_j^{-1} B_j^T \right)^T P_i \\ + Q_i - P_i B_i R_i^{-1} B_i^T P_i + A_p^T P_i A_p = 0. \tag{1.47}$$

Hence, the first part of Theorem 1.2 is proved. □

1.2.8 Stackelberg game

The concept of a Nash equilibrium solution is that there is no specific player controlling the decision process. However, there are other types of non-cooperative decision problems in which one participant has the ability to perform his/her strategy on other participants, and for such decision problems, a hierarchical equilibrium solution concept must be introduced. With regard to the work of H. von Stackelberg (1934), players who occupy a strong position in such a decision-making problem are called leaders, and other players involved in the decision-making of leaders are called followers. In two-player differential games, the existence of a hierarchy in decision-making implies that one of the players is in a position to determine his/her strategy ahead of time, announce it, and enforce it on the other players. Therefore, the Stackelberg solution is the only possible hierarchical equilibrium solution applicable in such decision-making problems, called Stackelberg games. A hierarchical equilibrium solution in Stackelberg games is generally called a Stackelberg strategy. Although there are many levels of hierarchy in the decision-making process, we limit the discussion here to two levels.

Let us consider the following stochastic system with linear quadratic cost functionals in the case of infinite-horizon:

$$dx(t) = [Ax(t) + B_0 u_0(t) + B_1 u_1(t)]dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.48a}$$

$$J_0(x_0,\ u_0,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t) Q_0 x(t) + u_0^T(t) R_{00} u_0(t) + u_1^T(t) R_{01} u_1(t)\right) dt\right], \tag{1.48b}$$

$$J_1(x_0,\ u_0,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t) Q_1 x(t) + u_0^T(t) R_{10} u_0(t) + u_1^T(t) R_{11} u_1(t)\right) dt\right] = 0, \tag{1.48c}$$

where A, B_0, B_1, A_p, $Q_i = Q_i^T \geq 0$, $R_{ii} = R_{ii}^T > 0$ and $R_{ij} = R_{ij}^T \geq 0,\ i \neq j$, for $i, j = 0, 1$ are the coefficient matrices of suitable dimensions. Assume that the Stackelberg game consists of two players. The initial state $x(0) = x_0$ is assumed to be a random variable with a covariance matrix $\mathbb{E}[x(0)x^T(0)] = I_n$.The subscript $i = 0$ represents the leader and $i = 1$ represents the follower. Assuming that both players adopt a closed loop strategy, the following definitions can be introduced.

Definition 1.8. *For any admissible strategy-set* $(u_0,\ u_1) \in \mathbb{U}$*, the strategy-set* $(u_0,\ u_1)$ *is called a Stackelberg strategy if the following conditions hold.*

$$J_0(x_0,\ u_0^*,\ u_1^*) \leq J_0(x_0,\ u_0,\ u_1^0(u_0)), \quad \forall u_0 \in \mathbb{R}^{m_0}, \tag{1.49}$$

where

$$J_1(x_0,\ u_0,\ u_1^0(u_0)) = \min_{u_1} J_1(x_0,\ u_0,\ u_1), \tag{1.50}$$

and

$$u_1^* = u_1^0(u_0^*). \tag{1.51}$$

It is well-known that the closed-loop Stackelberg strategies for the linear quadratic problems have the following form:

$$u_i = K_i x(t), \quad i = 0,\ 1. \tag{1.52}$$

Let us assume that $(A,\ B_i,\ A_p),\ i = 0,\ 1$ is stabilizable and $(A,\ A_p),\ i = 0,\ 1$ are exactly observable, then the following theorem can be derived.

Theorem 1.3. *[Mukaidani and Xu (2015a)] The strategy-set* (1.52) *constitutes the Stackelberg strategy only if the following cross-coupled stochsastic cross-coupled algebraic nonlinear matrix equations (ANMEs) have solutions* $P_1 \geq 0,\ M_0 \geq 0,\ N_1 > 0,\ N_0 > 0$ *and* K.

$$A_K^T P_1 + P_1 A_K + \hat{Q}_1 - P_1 B_1 R_{11}^{-1} B_1^T P_1 + A_p^T P_1 A_p = 0, \tag{1.53a}$$

$$A_{KF}^T P_0 + P_0 A_{KF} + \hat{Q}_0 + A_p^T P_0 A_p = 0, \tag{1.53b}$$

$$\begin{aligned} &M_1 A_K^T + A_K M_1 - M_1 P_1 B_1 R_{11}^{-1} B_1^T - B_1 R_{11}^{-1} B_1^T P_1 M_1 + A_p M_1 A_p^T - B_1 R_{11}^{-1} B_1^T P_0 M_0 \\ &\qquad - M_0 P_0 B_1 R_{11}^{-1} B_1^T + M_0 P_1 B_1 R_{11}^{-1} R_{01} R_{11}^{-1} B_1^T + B_1 R_{11}^{-1} R_{01} R_{11}^{-1} B_1^T P_1 M_0 = 0, \end{aligned} \tag{1.53c}$$

$$A_{KF} M_0 + M_0 A_{KF}^T + A_P M_0 A_p^T + I_n = 0, \tag{1.53d}$$

$$B_0^T (P_1 M_1 + P_0 M_0) + R_{01} K M_1 + R_{00} K M_0 = 0, \tag{1.53e}$$

where

$$F := -R_{11}^{-1} B_1^T P_1,$$

$$A_K := A + B_0 K,$$

$$
\begin{aligned}
&A_{KF} := A + B_0 K + B_1 F,\\
&\hat{Q}_1 := Q_1 + K^T R_{10} K,\\
&\hat{Q}_0 = Q_0 + F^T R_{01} F + K^T R_{00} K,
\end{aligned}
$$

Then the strategy-set $(u_0^*(t),\ u_1^*(t)) = (Kx(t),\ Fx(t))$ is the Stackelberg strategy.

Proof. For an arbitrary leader's strategy $u_0 = Kx(t)$, let us consider the follower's LQ closed-loop stochastic system as:

$$dx(t) = [(A + B_0 K)x(t) + B_1 u_1(t)]dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.54a}$$

$$J_1(x_0,\ Kx,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)(Q_1 + K^T R_{10} K)x(t) + u_1^T(t) R_{11} u_1(t)\right) dt\right]. \tag{1.54b}$$

Or, simply

$$dx(t) = [A_K x(t) + B_1 u_1(t)]dt + A_p x(t) dw(t), \quad x(0) = x^0, \tag{1.55a}$$

$$J_1(x_0,\ Kx,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)\hat{Q}_1 x(t) + u_1^T(t) R_{11} u_1(t)\right) dt\right], \tag{1.55b}$$

where

$$
\begin{aligned}
&A_K := A + B_0 K,\\
&\hat{Q}_1 := Q_1 + K^T R_{10} K.
\end{aligned}
$$

In fact (1.55) seems to be a standard LQ optimal control problem. So, if there exists a matrix $P_1^T = P_1 \geq 0$ that solves the following stochastic algebraic Riccati equation (SARE):

$$\mathbf{F}(P_1,\ K) := A_K^T P_1 + P_1 A_K + \hat{Q}_1 - P_1 B_1 R_{11}^{-1} B_1^T P_1 + A_p^T P_1 A_p = 0, \tag{1.56}$$

then the follower's state feedback control problem admits a solution,

$$u_1^0(u_0) = Fx(t) = F(K)x(t) = -R_{11}^{-1} B_1^T P_1 x(t). \tag{1.57}$$

It is seen that the SARE (1.56) is same as the SARE (1.53a). On the other hand, Leader cost J_0 can be obtained as,

$$
\begin{aligned}
J_0(x_0,\ u_0,\ u_1(u_0)) &= J_0(x_0,\ Kx,\ Fx)\\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)(Q_0 + F^T R_{01} F + K^T R_{00} K)x(t)\right) dt\right]\\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\infty x^T(t)\hat{Q}_0 x(t) dt\right] = Tr[P_0],
\end{aligned} \tag{1.58}
$$

with state equation

$$dx(t) = [A + B_0K + B_1F]x(t)dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{1.59}$$

where $P_0 = P_0^T \geq 0$ is the solution of the following SARE:

$$\mathbf{K}(P_1,\ P_0,\ K) := A_{KF}^T P_0 + P_0 A_{KF} + \hat{Q}_0 + A_p^T P_0 A_p = 0, \tag{1.60}$$

with

$$A_{KF} := A + B_0K + B_1F,$$
$$\hat{Q}_0 := Q_0 + F^T R_{01} F + K^T R_{00} K.$$

Therefore, SARE (1.53b) holds. Let us consider the Lagrangian $\mathscr{L}$ as follows:

$$\begin{aligned}\mathscr{L}(P_1,\ P_0,\ K) &= \mathbf{Tr}[P_0] + \mathbf{Tr}[M_1\mathbf{F}(P_1,\ K)] + \mathbf{Tr}[M_0\mathbf{K}(P_1,\ P_0,\ K)],\\ &= \mathbf{Tr}[P_0 + M_1(A_K^T P_1 + P_1 A_K + \hat{Q}_1 - P_1 B_1 R_{11}^{-1} B_1^T P_1 + A_p^T P_1 A_p)\\ &\quad + M_0(A_{KF}^T P_0 + P_0 A_{KF} + \hat{Q}_0 + A_p^T P_0 A_p)]\\ &= \mathbf{Tr}[P_0 + M_1 A_K^T P_1 + M_1 P_1 A_K + M_1 \hat{Q}_1 - M_1 P_1 B_1 R_{11}^{-1} B_1^T P_1 + M_1 A_p^T P_1 A_p\\ &\quad + M_0 A_{KF}^T P_0 + M_0 P_0 A_{KF} + M_0 \hat{Q} + M_0 A_p^T P_0 A_p)]\end{aligned} \tag{1.61}$$

where M_1 and M_0 are symmetric matrices of Lagrange multipliers. Using Lagrange multiplier technique for nonlinear matrix functions, the necessary conditions for minimizing $\mathbf{Tr}[P_0]$ are as follows:

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial P_1} &= M_1 A_K^T + A_K M_1 - M_1 P_1 B_1 R_{11}^{-1} B_1^T - B_1 R_{11}^{-1} B_1^T P_1 M_1 + A_p M_1 A_p^T - B_1 R_{11}^{-1} B_1^T P_0 M_0\\ &\quad - M_0 P_0 B_1 R_{11}^{-1} B_1^T + M_0 P_1 B_1 R_{11}^{-1} R_{01} R_{11}^{-1} B_1^T + B_1 R_{11}^{-1} R_{01} R_{11}^{-1} B_1^T P_1 M_0 = 0,\end{aligned} \tag{1.62a}$$

$$\frac{\partial \mathscr{L}}{\partial P_0} = A_{KF} M_0 + M_0 A_{KF}^T + A_P M_0 A_p^T + I_n = 0, \tag{1.62b}$$

$$\frac{1}{2}\frac{\partial \mathscr{L}}{\partial K} = B_0^T (P_1 M_1 + P_0 M_0) + R_{01} K M_1 + R_{00} K M_0 = 0. \tag{1.62c}$$

So, (1.53c)–(1.53e) are derived. Hence, Theorem 1.3 is proved. □

1.2.9 Incentive Stackelberg game

The incentive Stackelberg strategy is used to induce non-cooperative followers' virtual cooperation to achieve the optimal solution of the leader [Saksena and Cruz (1985)]. An incentive Stackelberg strategy is one where the leader achieves his/her team-optimal solution to the hierarchical game by using an incentive mechanism. The following two steps are the main elements of an incentive Stackelberg problem [Ho et al. (1982), Basar and Olsder (1980)].

(i) The leader determines a team-optimal strategy-set and announces it ahead of time.

(ii) Knowing the incentive, based on the leader's announced team-optimal strategy, each follower chooses a strategy so as to minimize his/her own cost.

It should be noted that no matter how the followers behave, the leader can achieve his/her own team-optimal equilibrium by using the corresponding incentive strategy-set. Incentive Stackelberg games apply to organizations with several players and with organizational objective functions that may not be the same as the members' objective functions. However, Chapter 3, Chapter 4 and Chapter 5 will discuss the disturbance attenuation incentive Stackelberg game in detail. Here we present only the basic formulation of the incentive Stackelberg game problem for one leader and one follower.

Consider a linear stochastic system governed by the Itô differential equation defined by

$$dx(t) = \big[Ax(t) + B_0 u_0(t) + B_1 u_1(t)\big]dt + A_p x(t)dw(t), \ \ x(0) = x^0, \tag{1.63}$$

where $x(t) \in \mathbb{R}^n$ represents the state vector; $u_0(t) \in \mathbb{R}^{m_0}$ represents the leader's control input for the follower; $u_1(t) \in \mathbb{R}^{m_1}$ represents the follower 's control input; $w(t) \in \mathbb{R}$ represents a one-dimensional standard Wiener process defined in the filtered probability space $(\Omega,\ \mathscr{F},\ \mathscr{P},\ \mathscr{F}_t)$ with $\mathscr{F}_t = \sigma\{w(s)\ :\ 0 \le s \le t\}$ [Chen and Zhang (2004)].

Cost functionals of the leader is given by

$$J_0\,(x_0,\ u_0,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^{\infty}\left\{x^T(t)Q_0x(t) + u_0^T(t)R_{00}u_0(t) + u_1^T(t)R_{01}cu_1(t)\right\}dt\right], \tag{1.64}$$

where $Q_0 = Q_0^T \ge 0$, $R_{00} = R_{00}^T > 0$, $R_{01} = R_{01}^T \ge 0$.

Cost functionals of the follower is given by

$$J_1\,(x_0,\ u_0,\ u_1) := \frac{1}{2}\mathbb{E}\left[\int_0^{\infty}\left\{x^T(t)Q_1x(t) + \left[u_1^T(t)R_{10}u_1(t) + u_1^T(t)R_{11}u_1(t)\right]\right\}dt\right], \tag{1.65}$$

where $Q_1 = Q_1^T \ge 0$, $R_{11} = R_{11}^T > 0$ and $R_{10} = R_{10}^T \ge 0$. For a two-level incentive Stackelberg game, leader announce the following incentive strategy to the follower in ahead of time:

$$u_0(t) = \Lambda x(t) + \Xi u_1(t), \tag{1.66}$$

where the parameters Λ and Ξ are to be determined associated with the optimal strategy $u_1(t)$ of the follower.

First, the leadear's team-optimal solution $u_c^*(t)$ is investigated. By composing the stochastic system (1.63), the following centralized systems can be obtained.

$$dx(t) = [Ax(t) + B_c u_c(t)]\,dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{1.67}$$

where

$$B_c := \begin{bmatrix} B_0 & B_1 \end{bmatrix},$$
$$u_c := \mathbf{col}\begin{bmatrix} u_0 & u_1 \end{bmatrix}.$$

Furthermore, the cost functional (1.64) can be modified as

$$J_0\left(u_c(t)\right) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)Q_0x(t) + u_c^T(t)R_cu_c(t)\right\} dt\right], \tag{1.68}$$

where

$$R_c := \textbf{block diag}\begin{bmatrix} R_{00} & R_{01} \end{bmatrix}.$$

Therefore, using standard LQ theory, suppose that the leader team-optimal state feedback strategy is given by

$$u_c^*(t) = K_cx(t) = -R_c^{-1}B_c^TP_cx(t), \tag{1.69}$$

where

$$K_c = \begin{bmatrix} K_{c0} \\ K_{c1} \end{bmatrix} = \begin{bmatrix} -R_{00}^{-1}B_0^TP_c \\ -R_{01}^{-1}B_1^TP_c \end{bmatrix}. \tag{1.70}$$

Then there exists a matrix $P_c^T = P_c \geq 0$ that solves the following SARE:

$$P_cA + A^TP_c + Q_0 - P_cB_cR_c^{-1}B_c^TP_c + A_p^TP_cA_p = 0. \tag{1.71}$$

It should be noted that the relation between Λ and Ξ can be derived from (1.66) as

$$-R_{00}^{-1}B_0^TP_c = \Lambda - \Xi R_{01}^{-1}B_1^TP_c, \tag{1.72}$$

or,

$$\Lambda = -R_{00}^{-1}B_0^TP_c + \Xi R_{01}^{-1}B_1^TP_c. \tag{1.73}$$

So, the leader's incentive Stackelberg strategy can be determined by

$$\begin{aligned} u_0(t) &= [K_{c0} - \Xi K_{c1}]\,x(t) + \Xi u_1(t), \\ &= K_{c0}x(t) + \Xi\left[u_1 - K_{c1}x(t)\right]. \end{aligned} \tag{1.74}$$

To determine Ξ which satisfy (1.66), let us consider the following optimization problem for the follower. To establish the follower's according optimal strategy regarding leader's incentive Stackelberg strategy (1.66), we can change (1.63) as follows,

$$\begin{aligned} dx(t) &= \left[Ax(t) + B_0u_0(t) + B_1u_1(t)\right]dt + A_px(t)dw(t),\ \ x(0) = x^0, \\ &= \left[Ax(t) + B_0(\Lambda x(t) + \Xi u_1(t)) + B_1u_1(t)\right]dt + A_px(t)dw(t),\ \ x(0) = x^0, \end{aligned}$$

$$= \left[\hat{A}x(t) + \hat{B}u_1(t)u_1(t)\right]dt + A_p x(t)dw(t), \;\; x(0) = x^0. \tag{1.75}$$

where

$$\hat{A} := A + B_0\Lambda,$$
$$\hat{B} := B_1 + B_0\Xi.$$

The cost functional of the follower can be written as

$$\begin{aligned} J_1(x_0,\, u_0,\, u_1) =& \frac{1}{2}\mathbb{E}\Big[\int_0^{\infty} \big[x^T(t)Q_1x(t) + (\Lambda x(t) + \Xi u_1(t))^T R_{10}(\Lambda x(t) + \Xi u_1(t)) \\ & + u_1^T(t)R_{11}u_1(t)\big]dt\Big], \\ =& \frac{1}{2}\mathbb{E}\Bigg[\int_0^{\infty} \big\{x^T(t)\hat{Q}x(t) + \big[[\Lambda x(t) + \Xi u_1(t)]^T R_{10}[\Lambda x(t) + \Xi u_1(t)] \\ & + u_1^T(t)R_{11}u_1(t)\big]\big\}dt\Bigg], \end{aligned} \tag{1.76}$$

or, equivalently,

$$J_1(x_0,\, u_1) = \frac{1}{2}\mathbb{E}\left[\int_0^{\infty}\left\{x^T(t)\hat{Q}x(t) + 2x^T(t)\hat{S}u_1(t) + u_1^T(t)\hat{R}u_1(t)\right\}dt\right], \tag{1.77}$$

where

$$\hat{Q} := Q_0 + \Lambda^T R_{10}\Lambda,$$
$$\hat{R} := R_{11} + \Xi^T R_{10}\Xi,$$
$$\hat{S} := \Lambda^T R_{10}\Xi.$$

It should be noted that there exists a cross-coupling term $2x^T(t)\hat{S}u_1(t)$ in the cost functional (1.77). By using the technique similar to the one used in the stochastic optimal control problem [Chen and Zhang (2004)], follower's optimal strategy $u_1^{\dagger}(t) = K_1^{\dagger}x(t)$ can be obtained, where

$$K_{Fi}^{\dagger} = -\hat{R}^{-1}(P_1\hat{B} + \hat{S})^T, \tag{1.78}$$

and P_1 is the symmetric non-negative solution of the following SARE:

$$P_1\hat{A} + \hat{A}^T P_1 + A_p^T P_1 A_p - (P_1\hat{B} + \hat{S})\hat{R}^{-1}(P_1\hat{B} + \hat{S})^T + \hat{Q} = 0. \tag{1.79}$$

Furthermore, to keep the optimality of the entire system unchanged, followers have to determine Ξ satisfying the following equivalence relation.

$$K_{c1}^{*} \equiv K_1^{\dagger},$$

which can establish from (1.70) and (1.78) as follows:

$$-R_{01}^{-1}B_1^T P_c = -\hat{R}^{-1}(P_1\hat{B}+\hat{S})^T, \tag{1.80}$$

or,

$$[R_{11}+\Xi^T R_{10}\Xi]R_{01}^{-1}B_1^T P_c = B_1^T P_1 + \Xi^T B_0^T P_1 + \Xi^T R_{10}\Lambda. \tag{1.81}$$

By using relation (1.73) for Λ, we get

$$\begin{aligned} &R_{11}R_{01}^{-1}B_1^T P_c + \Xi^T R_{10}\Xi R_{01}^{-1}B_1^T P_c = \\ &\quad B_1^T P_1 + \Xi^T B_0^T P_1 - \Xi^T R_{10}R_{01}^{-1}B_0^T P_c + \Xi^T R_{10}\Xi R_{01}^{-1}B_1^T P_c. \end{aligned} \tag{1.82}$$

Canceling the term $\Xi^T R_{10}\Xi R_{01}^{-1}B_1^T P_c$ from both sides of (1.82) we get

$$R_{11}R_{01}^{-1}B_1^T P_c = B_1^T P_1 + \Xi^T B_0^T P_1 - \Xi^T R_{10}R_{01}^{-1}B_0^T P_c, \tag{1.83}$$

and after simplification, the following MAEs can be found:

$$\Xi^T(B_0^T P_1 - R_{10}R_{01}^{-1}B_0^T P_c) = R_{11}R_{01}^{-1}B_1^T P_c - B_1^T P_1. \tag{1.84}$$

Remark 1.2. *It should be noted that the incentive parameter Ξ can be uniquely determined if and only if $(B_0^T P_1 - R_{10}R_{01}^{-1}B_0^T P_c)$ is non-singular.*

Theorem 1.4. *Suppose that the SARE in* (1.71)*, SARE* (1.79) *and the MAEs* (1.83) *have solutions. Then the strategy-set* (1.66) *under* (1.70) *and* (1.78) *constitutes the two-level incentive Stackelberg strategies with H_∞ constraint.*

1.3 Research survey

The Stackelberg leadership model is a hierarchical strategy involving the first movement of the leader and then the consequent movement of the followers. The term Stackelberg was used after the German economist Heinrich Freiherr von Stackelberg, who introduced this idea in his article 'Market structure and equilibrium (Marktform und Gleichgewicht)' in 1934. The properties of Stackelberg games for two players have been extensively studied in [Starr and Ho (1969)]. Subsequently, this two-player static game was extended to a dynamic game with different information patterns [Chen and Cruz (1972), Simaan et al. (1973)]. Among the information patterns, closed-loop Stackelberg strategies with applications were attracting considerable research interest as – LQ problems [Medanic (1978), Basar and Selbuz (1979), Tolwinski (1981)].

The idea of team-optimal solutions opens new directions for closed-loop Stackelberg strategies. In [Basar and Olsder (1980)], the necessary and sufficient conditions for both finite- and infinite-horizon closed-loop feedback solutions were derived for a team problem in which all players optimized a leader's cost functional jointly. Furthermore, [Salman and Cruz (1983)] derived team-optimal closed-loop Stackelberg strategies for systems with slow and fast modes.

An incentive Stackelberg strategy is one where the leader achieves his/her team-optimal solution to the hierarchical game by using an incentive mechanism. Through the last four decades, incentive Stackelberg games are the growing interest in research (see, e.g., [Basar and Olsder (1980), Ho et al. (1982), Ishida and Shimemura (1983), Zheng et al. (1984), Mizukami and Wu (1987), Mizukami and Wu (1988)], and references therein). The purpose of the incentive mechanism is to induce virtual cooperation in non-cooperative followers so that optimal system performance (reflected in the leader's objective function) is achieved through hierarchical decision-making [Saksena and Cruz (1985)]. A two-person, continuous-time, linear differential game problems under quadratic cost functions for both players were derived in [Basar and Olsder (1980)]. In that paper, the authors treated both the finite- and the infinite-horizon cases. Unlike the discrete-time version of [Basar and Selbuz (1979)], which article used a closed-loop team-optimal strategy in a continuous-time differential game. In the recent years, there were many papers and works dealing with the incentive Stackelberg strategy. In [Zheng and Basar (1982)], the existence and derivation of the affine-excitation Stackelberg strategy were studied by using the geometric method. In [Zheng et al. (1984)], the closed-loop Stackelberg strategy and incentive policy in the dynamic decision problem were widely discussed. With several control problems, dynamic games for both continuous- and discrete-time systems have been extensively studied (see e.g. [Başar and Olsder (1999)] and references therein).

In [Mizukami and Wu (1987), Mizukami and Wu (1988)], incentive Stackelberg strategies were derived for LQ differential games, where the two leaders and one follower to the first paper and one leader and two followers to the second paper were considered. In recent years, incentive Stackelberg games with robust control theory have been studied for discrete-time linear systems with a deterministic disturbance in [Ahmed and Mukaidani (2016), Mukaidani et al. (2017c)]. In both articles, one leader and multiple non-cooperative followers were considered. A similar structure was adapted in [Ahmed et al. (2017a), Mukaidani et al. (2017d), Mukaidani and Xu (2018)] for stochastic case with H_∞ constraint. On the other hand, continuous-time stochastic systems were investigated for an infinite-horizon incentive Stackelberg game in [Mukaidani (2016), Ahmed et al. (2017b)], where multiple leaders and multiple followers were considered. In

[Mukaidani et al. (2017b)], incentive Stackelberg strategy with multiple leaders and multiple followers were considered for the stochastic Markov jumping control problem. It should be noted that [Ahmed et al. (2017b)] discussed two-level hierarchical stochastic games with M leaders and N followers with H_∞ constraint.

On the other hand, a linear parameter-varying (LPV) system was introduced in [Shamma (1988)] for analyzing the "gain-scheduling" problem. Gain scheduling involves implementing a family of linear controllers such that controller coefficients (gains) are changed (scheduled) based on the current values of the scheduled variables. In short, gain scheduling is a control design approach that builds a nonlinear controller for a nonlinear plant by patching a set of linear controllers.

1.4 Objectives and outlines

The primary objective of this research is to investigate the incentive Stackelberg strategy for various dynamic systems. The incentive Stackelberg games are examined for the systems including external disturbances which are attenuated under the H_∞ constraint. Solving game problems of multiple players with disturbance terms is more difficult while stochastic systems are observed. In the incentive Stackelberg games, players are divided into two groups, leaders, and followers. Therefore, games with several constraints cause complicated problems including accuracy of results. In other words, the incentive Stackelberg game with such a complex pattern is a new approach in the two-level hierarchy. For this approach, it is necessary to face the computational complexity for solving algebraic Riccati equations (AREs). Besides, preliminary research and synthesis of the LPV system is provided for multiple decision makers. We aim to better understand to implement our current idea for LPV easily in further.

This thesis attempts to consider the incentive Stackelberg game for both the discrete- and continuous-time framework, both deterministic and stochastic systems are investigated. In the case of continuous-time and LPV systems, only stochastic structures are included in this thesis. For all schemes, linear quadratic optimal control is investigated only for simplicity. However, the investigation of nonlinear problems is our future research. Therefore, after discussing the basic problem of stochastic LQ systems, four major chapters can be viewed afterward.

This thesis consists of seven chapters. In Chapter 1, the research background, motivation, research survey, objectives, and outlines of the thesis are described. Some basic definitions and preliminary results are also introduced in this chapter. Chapter 2 of the thesis summarizes some of the preliminary mathematical problems for this study. This chapter is an overview of the stochastic LQ systems. Several problems based on discrete-time and continuous-time stochastic optimal control are solved. Stochastic dynamic programming (DP) and Itô's lemma are outlined. In particular, we discuss the concept of the solution of multi-player games considering Nash equilibrium and Pareto optimality. The exogenous disturbance problem and its attenuation of the H_∞ constraint are presented and will play an important role in later developments.

In Chapter 3, the incentive Stackelberg game for a discrete-time deterministic system is considered. It explains two levels of hierarchy with one leader and multiple followers. Followers are supposed to act non-cooperatively. Exogenous disturbance also exists in the system and is attenuated under the H_∞ constraint. In this work, the team-optimum solution of the leader is formulated based on the result of [Zhang et al. (2007)]. By simulating a set of cross-coupled backward difference Riccati equations (CCBDREs), the leader's team optimal solution can be found. The H_∞ constraint for disturbance attenuation is taken into account at the same time. On the other hand, the optimal strategy of the followers based on the Nash equilibrium guarantees the leader's team optimal solution. The expression is also derived for the infinite-horizon case. An algorithm for solving the set of cross-coupling algebra Riccati equations (CCARE) is developed. A numerical example shows the efficiency of the proposed method.

Chapter 4 investigates the incentive Stackelberg game for discrete-time stochastic systems. The structure of this game is very similar to that described in Chapter 3. It is a single leader and multiple non-cooperative followers with an exogenous disturbance attenuating under the H_∞ constraint in the 2-level hierarchy. Therefore, Chapter 4 can be viewed as the stochastic version of the deterministic game described in Chapter 3. In this chapter, we determine the leader's incentive Stackelberg strategies according to the result of [Zhang et al. (2007)] of the stochastic discrete-time system with disturbance. The information pattern of this problem question is as follows. The leader can access all the values of the follower's decision at each stage of the process. An incentive mechanism for leading non-cooperative followers is virtual cooperation to achieve system goals. To solve the problem in the case of finite- and infinite-horizon, a set of cross-coupled *stochastic backward difference Riccati equation* (SBDRE) and *stochastic matrix-valued difference equations* (SMVDEs) are derived, correspondingly. Furthermore, the Nash equilibrium of the followers guarantees the leader's team optimal solution. A Lyapunov-based recursive algorithm has also been designed to reduce the computational complexity. Academic and practical numerical examples guarantee the efficiency of the proposed method.

In Chapter 5, the continuous-time incentive Stackelberg games for multiple leaders and multiple followers are investigated. The external disturbance is included with the system, as usual. To explain this kind of game, consider that the M leader and the N follower belong to two groups. By adjusting the values of M and N, we can form any suitable hierarchical game. That is why we call it a generic construction. The information pattern of the game is more complex than before we followed. Each leader must achieve the Nash equilibrium and use the H_∞ constraint to reduce the external disturbance. Each leader will individually announce the incentive Stackerberg strategy for each follower. Each follower employs a leader incentive mechanism that follows the Nash equilibrium in a follower group. Leaders and followers do not cooperate with their group. In this chapter, the Nash equilibrium of the leader under the H_∞ constraint is derived based on the infinite-horizontal stochastic H_2/H_∞ control problem of [Chen and Zhang (2004)]. We can get the strategy-set by solving a set of *cross-coupled stochastic algebraic Riccati equations* (CCSAREs) and *matrix algebraic equations* (MAEs). The leaders achieve the Nash equilibrium by attenuating disturbance under the H_∞ constraint. Simultaneously, the followers achieve the Nash equilibrium regarding the leaders' incentives with the worst-case disturbance. The solution can be easily found using Lyapunov-based iterations. To illustrate our findings, we present a simple numerical example.

Chapter 6 discusses the Pareto optimal strategy for the stochastic LPV system with multiple players. In the dynamic game of uncertain stochastic systems, multiple participants can be used for more realistic plants. The system includes disturbances that are attenuated under the H_∞ constraint. This section can be seen as an extension of [Mukaidani (2017a)]. This is because the fixed gain controller is also considered here to understand the practical implementation. In this chapter, we design a method for Pareto optimal solution that satisfies the H_∞ norm condition. We redesigned the stochastic bounded real lemma [Ku and Wu (2015)] and the linear quadratic control [Rotondo (2015)] to find the solution. Solvability conditions are established using linear matrix inequalities (LMIs). For multiple players, a Pareto optimal strategy-set is designed. The Pareto optimal strategy-set can be found by solving a set of cross-coupling matrix inequalities (CCMI). A numerical example is provided to demonstrate the effectiveness of the proposed model of the LPV system. However, for stochastic LPV systems, the H_∞ constraint incentive Stackelberg game is not investigated in this chapter. This will be our future research. It should be noted that some basic results on the LMI problems are presented in Appendix A as a preliminary study of this chapter.

Finally, in Chapter 7, the thesis is concluded with some motivating guidelines for future research. It should be noted that Appendix B contains some of the basic results of output feedback control as a preliminary study of future research.

Chapter 2

Basic Problems of Stochastic Linear Quadratic (LQ) Systems

Optimization problems can be divided into two groups, static and dynamic. An optimization problem that does not change over time is called a static optimization problem. Many scientific and business applications require the control of systems developed over time, called dynamic systems. Since time can vary discretely and continuously, dynamic systems are separated into discrete- and continuous-time systems. This study plans to deal with both types of systems, with the emphasis on continuous-time systems. Optimal control of the dynamic system addresses the finding of control actions achieved in an optimal manner ensuring the stability of the control. There are various types of stability in describing solutions to difference (for discrete-time) or differential (for continuous-time) equations in dynamical systems. The most important type is that solution stability approaches the equilibrium point discussed in the Lyapunov stability theory. An optimal control problem consists of a cost functional and a set of difference or differential equations describing the trajectories of the control variables that minimize the cost functional.

In this research, we focus only on the linear control system composed of linear difference or differential equations. There are several reasons for choosing linear optimal control instead of general optimal control. For example, many engineering problems are linear before adding controllers. It is easy to physically implement, takes less time to calculate, applies to small signal based nonlinear systems, and computation algorithm proposes nonlinear optimization design. The specialty of linear control is that the plant to be controlled as well as the control unit that produces optimal control are assumed to be linear. The linear optimal controllers are attained by operating with quadratic performance indices. Such pattern of linear control that minimizes the sum (for discrete-time) or integral (for continuous-time) of the quadratic function assessed by the control and state variables is called linear quadratic (LQ) optimal control. LQ optimal control initiated by Kalman

[Kalman (1960)] plays a major role in control theory. Many researchers have conducted extensive studies on deterministic LQ problems [Anderson and Moore (1989), Lewis (1986)]. In [Wonham (1968)], Wonham introduced the stochastic LQ optimal control governed by the Ito differential equations. Systems perturbed by Gaussian white noise are called linear quadratic Gaussian control problems and have become the most popular in control theory research [Athans (1971)].

The main idea of LQ control design is to minimize the quadratic cost functional, $\int_0^\infty (x^T Qx + u^T Ru)dt$. It turns out that regardless of the values of Q and R, the cost functional has a unique minimum that can be obtained by solving the Algebraic Riccati Equation. The parameters Q and R can be used as design parameters to penalize the state variables and the control signals. The larger these values are, the more we penalize these signals. Basically, choosing a large value for R means we try to stabilize the system with less (weighted) energy. This is usually called an expensive control strategy. On the other hand, choosing a small value for R means we do not want to penalize the control signal (cheap control strategy). Similarly, if we choose a large value for Q means we try to stabilize the system with the least possible changes in the states and large Q implies less concern about the changes in the states. Since there is a trade-off between the two, we may want to keep Q as I (identity matrix) and only alter R. We can choose a large R, if there is a limit on the control output signal (for instance, if large control signals introduce sensor noise or cause actuator's saturation), and choose a small R if having a large control signal is not a problem for the system.

Indeed, with random choice of Q and R matrices, the optimal regulators do not provide good set point tracking performance. Conventionally, control engineers often select the weighting matrices based on trial and error approach, which not only makes the design tedious but also provides a non optimized response. One more methodology adopted in the design of optimal controller is that the initial values of weighting matrices could be chosen as $Q = C^T C$ and $R = B^T B$, where C comes from the controlled output $y = Cx$ and B is the coefficient matrix of the control input (u); and after the initial trial, if the performance is not satisfactory these weights can be altered again to get the desired response. However, this approach once again makes use of trial and error method.

2.1 Discrete-time stochastic optimal control problems

In the framework of discrete-time, the decision maker observes state variables for each time period. The objective is to optimize the sum of expected values of the objective function over the entire period. New observations are made in each time period, and the control variables are optimally adjusted. The optimal solution for the current time can be found by iterating the matrix Riccati equation from the last period back to the current period. The discrete-time stochastic linear quadratic optimal control problem can be expressed as follows:

$$J(x_0,\, u) := \min \frac{1}{2}\mathbb{E}\left[x^T(T_f)Q(T_f)x(T_f) + \sum_{k=0}^{T_f-1}[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k)]\right], \tag{2.1}$$

where $Q(k) = Q^T(k) \geq 0$; $R(k) = R^T(k) > 0$; $\mathbb{E}[\cdot]$ is the expected operator conditional on x_0; superscript T indicates the transpose of a matrix and $0 < T_f < \infty$ is the time range. The state equation is defined as follows:

$$x(k+1) = A(k)x(k) + B(k)u(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \tag{2.2}$$

where $x(t) \in \mathbb{R}^n$ is the state vector; $u(t) \in \mathbb{R}^{n_u}$ is the control input; $w(t) \in \mathbb{R}$ is a one-dimensional wiener process; and $A(k)$, $B(k)$, $A_p(k)$ are the coefficient matrices of suitable dimensions.

Theorem 2.1. *If the following stochastic backward difference Riccati equation (SBDRE) has a solution $P(k) > 0$:*

$$\begin{aligned} P(k) =& A^T(k)P(k+1)A(k) + A_p^T(k)P(k+1)A_p(k) - A^T(k)P(k+1)B(k) \\ & \times [R(k) + B^T(k)P(k+1)B(k)]^{-1}B^T(k)P(k+1)A(k) + Q(k),\; P(T_f+1) = 0, \end{aligned} \tag{2.3}$$

then, the discrete-time stochastic system (2.2) *with cost functional* (2.1) *have the following optimal state feedback control:*

$$u^*(k) = -[R(k) + B^T(k)P(k+1)B(k)]^{-1}B^T(k)P(k+1)A(k)x(k). \tag{2.4}$$

Proof. According to dynamic programming algorithm we can write discrete-time stochastic quadratic Hamilton-Jacobi-Bellman (HJB) as follows:

$$\begin{aligned} V(k) &= \min_{u(k)} \frac{1}{2}\mathbb{E}\left[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k) + V(k+1)\right] \\ &= \min_{u(k)} \frac{1}{2}\mathbb{E}\Big[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k) + x^T(k+1)P(k+1)x(k+1) + 2v(k+1)\Big], \end{aligned} \tag{2.5}$$

where $V(k)$ is a quadratic function with a stochastic increment as follows:

$$V(k) = \frac{1}{2}x^T(k)P(k)x(k) + v(k) \tag{2.6a}$$

$$v(k+1) = \frac{1}{2}\mathbf{Tr}[P(k+1)A_p(k)x(k)x^T(k)A_p^T(k)] + v(k), \quad v(T_f) = 0, \tag{2.6b}$$

with the symmetric positive definite matrix P.

Using equation (2.6b), the following result can be found form (2.5):

$$\begin{aligned}
V(k) =& \min_{u(k)} \frac{1}{2}\mathbb{E}\Big[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k) \\
&+ [A(k)x(k)+B(k)u(k)]^T P(k+1)[A(k)x(k)+B(k)u(k)] \\
&+ \mathbf{Tr}[P(k+1)A_p(k)x(k)x^T(k)A_p^T(k)] + 2v(k)\Big], \\
=& \min_{u(k)} \frac{1}{2}\mathbb{E}\Big[x^T(k)[Q(k)+A^T(k)P(k+1)A(k)+A_p^T(k)P(k+1)A_p(k)]x(k) \\
&+ u^T(k)[R(k)+B^T(k)P(k+1)B(k)]u(k) + 2x^T(k)A^T(k)P(k+1)B(k)u(k)\Big] + v(k),
\end{aligned} \tag{2.7}$$

To minimize the right hand side of (2.7) with respect to the control input $u(k)$ we obtain the following state feedback optimal control scheme:

$$u^*(k) = -[R(k)+B^T(k)P(k+1)B(k)]^{-1}B^T(k)P(k+1)A(k)x(k), \tag{2.8}$$

Comparing right hand sides of (2.6a) and (2.7), the following stochastic backward difference Riccati equation (SBDRE) can be derived using (2.8):

$$\begin{aligned}
P(k) =& A^T(k)P(k+1)A(k) + A_p^T(k)P(k+1)A_p(k) - A^T(k)P(k+1)B(k) \\
&\times [R(k)+B^T(k)P(k+1)B(k)]^{-1}B^T(k)P(k+1)A(k) + Q(k), \ P(T_f+1) = 0.
\end{aligned} \tag{2.9}$$

□

For infinite-horizon case, the state equation and the cost functional have the following form:

$$x(k+1) = Ax(k) + Bu(k) + A_p x(k)w(k), \quad x(0) = x_0, \tag{2.10a}$$

$$J(x_0,\ u) := \min \frac{1}{2}\mathbb{E}\left[\sum_{k=0}^{\infty}[x^T(k)Qx(k) + u^T(k)Ru(k)]\right], \tag{2.10b}$$

where $Q = Q^T \geq 0,\ R = R^T > 0$. It should be noted that for the infinite-horizon case, the coefficient matrices are considered to be constant matrices of appropriate dimensions.

Lemma 2.1. *Suppose that there exists a symmetric constant matrix $P > 0$ that solves the following SARE of the system* (2.10)*:*

$$P = A^T PA + A_p^T PA_p - A^T PB[R+B^T PB]^{-1}B^T PA + Q. \tag{2.11}$$

then the optimal control problem admits a state feedback solution,

$$u(k) = -[R+B^T PB]^{-1}B^T PAx(k), \tag{2.12}$$

Proof. Using optimal control $u^*(k) = Kx(k)$, the state feedback system (2.10) can be written as

$$x(k+1) = (A+BK)x(k) + A_p x(k)w(k), \ \ x(0) = x_0, \tag{2.13}$$

with cost functional

$$J(x_0) := \min \frac{1}{2}\mathbb{E}\left[\sum_{k=0}^{\infty} x^T(k)[Q+K^T RK]x(k)\right]. \tag{2.14}$$

Suppose that there exits a symmetric positive definite matrix P such that the SARE (2.11) holds for all admissible control inputs. Let us define the Lyapunov candidate function

$$\mathbb{E}[V(x(k))] = \mathbb{E}[x^T(k)Px(k)], \tag{2.15}$$

where $V(x(k) > 0$ for all $x(k) \neq 0$.

The difference between corresponding trajectory of the system (2.10) is given by

$$\begin{aligned} \mathbb{E}[\Delta V(x(k))] &= \mathbb{E}[V(x(k+1)) - V(x(k))] \\ &= \mathbb{E}[x^T(k+1)Px(k+1) - x^T(k)Px(k)] \\ &= \mathbb{E}[x^T(k)(A+BK)^T P(A+BK)x(k)] \\ &\quad + \mathbb{E}[x^T(k)A_p^T PA_p x(k)] - \mathbb{E}[x^T(k)Px(k)] \\ &= \mathbb{E}[x^T(k)[(A+BK)^T P(A+BK) + A_p^T PA_p - P]x(k), \end{aligned} \tag{2.16}$$

which is stable if $\mathbb{E}[\Delta V(x(k))] < 0$. Then, we can form the discrete-time Lyapunov stabilizable equation [Zhang et al. (2008)] as follows:

$$(A+BK)^T P(A+BK) + A_p^T PA_p - P = -(Q+K^T RK) \tag{2.17}$$

Substituting the value of $K = -[R+B^T PB]^{-1}B^T PA$ to equation (2.17) and simplifying, we can get the following SARE:

$$P = A^T PA + A_p^T PA_p - A^T PB[R+B^T PB]^{-1}B^T PA + Q. \tag{2.18}$$

Hence, Lemma 2.1 is proved. □

2.2 Continuous-time stochastic optimal control problems

Consider the following continuous-time stochastic linear quadratic optimal control problem:

$$dx(t) = [A(t)x(t) + B(t)u(t)]dt + A_p(t)x(t)dw(t), \quad x(0) = x^0, \tag{2.19a}$$

$$J(x_0,\ u) := \frac{1}{2}\mathbb{E}\left[x^T(T_f)Q(T_f)x(T_f) + \int_0^{T_f} \left(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\right) dt\right], \quad (2.19b)$$

where $x(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^n)$ is the state vector; $u(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ is the control input; $w(t) \in \mathbb{R}$ is a one-dimensional wiener process; $A(t)$, $B(t)$, $A_p(t)$, $Q(t) = Q^T(t) \geq 0$, $R(t) = R^T(t) > 0$ are the coefficient matrices of suitable dimensions; $\mathbb{E}[\cdot]$ is the expected operator conditional on x_0; $0 < T_f < \infty$ is the time range; $\mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^\ell)$ denotes the space of nonanticipative stochastic processes. In order to solve the above-mentioned optimal control problem, the following theorem can be obtained.

Theorem 2.2. *For the stochastic optimal control problem* (2.19)*, suppose that the following stochastic Riccati differential equation (SRDE) has the solution* $P^T(t) = P(t) \geq 0$*:*

$$\begin{aligned} -\dot{P}(t) =& A^T(t)P(t) + P(t)A(t) + A_p^T(t)P(t)A_p(t) + Q(t) \\ & - P(t)B(t)R^{-1}(t)B^T(t)P(t), \quad P(T_f) = Q(T_f), \end{aligned} \quad (2.20)$$

then the optimal control problem admits a state feedback solution,

$$u^*(t) = -R^{-1}(t)B^T(t)P(t)x(t). \quad (2.21)$$

Proof. In order to prove the Theorem 2.2, the stochastic dynamic programming (SDP) method and the stochastic maximum principle can be considered. The following two sections derive the Theorem 2.2 as two different techniques.

2.2.1 Stochastic dynamic programming (SDP)

We define the finite-horizon value function,

$$v(t,x) = \frac{1}{2} \min_{u(t) \in U} \mathbb{E}\left[\int_t^{T_f} \{x^T(s)Q(s)x(s) + u^T(s)R(s)u(s)\} ds\right]. \quad (2.22)$$

It will satisfy the stochastic Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned} -v_t = \min_{u(t) \in U} [& \frac{1}{2}(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)) + v_x^T(A(t)x(t) + B(t)u(t)) \\ & + \frac{1}{2}(A_p(t)x(t))^T(A_p(t)x(t))v_{xx}], \end{aligned} \quad (2.23)$$

with boundary condition: $v(T_f, x) = \frac{1}{2}x^T(T_f)Q(T_f)x(T_f)$.

To minimize the right hand side of (2.23) with respect to the control input $u(t)$ we get,

$$R(t)u^*(t) + B^T(t)v_x = 0, \quad (2.24)$$

or,

$$u^*(t) = -R^{-1}(t)B^T(t)v_x. \tag{2.25}$$

Now, if we insert this optimal state feedback control input $u^*(t)$ into equation (2.23) we get,

$$\begin{aligned}
-v_t =& \frac{1}{2}(x^T(t)Q(t)x(t) + v_x^T B(t)R^{-1}(t)B^T(t)v_x) + v_x^T(A(t)x(t) - B(t)R^{-1}(t)B^T(t)v_x) \\
&+ \frac{1}{2}x^T(t)A_p^T(t)A_p(t)x(t)v_{xx}, \\
=& \frac{1}{2}(x^T(t)Q(t)x(t) - v_x^T B(t)R^{-1}(t)B^T(t)v_x) + v_x^T A(t)x(t) + \frac{1}{2}x^T(t)A_p^T(t)A_p(t)x(t)v_{xx},
\end{aligned} \tag{2.26}$$

with boundary condition:

$$v(T_f, x) = \frac{1}{2}x^T(T_f)Q(T_f)x(T_f).$$

Now let

$$v(t,x) = \frac{1}{2}x^T(t)P(t)x(t), \tag{2.27}$$

where $P(t)$ is a symmetric positive semidefinite matrix. Therefore,

$$v_x = P(t)x(t),$$

and

$$v_{xx} = P(t).$$

So, from (2.26) we get,

$$\begin{aligned}
-\frac{1}{2}x^T(t)\dot{P}(t)x(t) =& \frac{1}{2}(x^T(t)Q(t)x(t) - x^T(t)P(t)B(t)R^{-1}(t)B^T(t)P(t)x(t)) \\
&+ \underbrace{x^T(t)P(t)A(t)x(t)}_{\frac{1}{2}x^T(t)(P(t)A(t)+A^T(t)P(t))x(t)} + \frac{1}{2}x^T(t)A_p^T(t)P(t)A_p(t)x(t),
\end{aligned} \tag{2.28}$$

with boundary condition: $P(T_f) = Q(T_f)$. Simply,

$$\begin{aligned}
-\dot{P}(t) =& A^T(t)P(t) + P(t)A(t) + A_p^T(t)P(t)A_p(t) + Q(t) \\
&- P(t)B(t)R^{-1}(t)B^T(t)P(t), \quad P(T_f) = Q(T_f).
\end{aligned} \tag{2.29}$$

Moreover, substituting $v_x = P(t)x(t)$ into equation (2.25), we can obtain the state feedback optimal control input,

$$u^*(t) = -R^{-1}(t)B^T(t)P(t)x(t). \tag{2.30}$$

2.2.2 Stochastic maximum principle

Recall the continuous-time stochastic system (2.19a) with linear quadratic cost (2.19b). In order to find the solution of this optimal control problem by stochastic maximum principle [Peng (1990)], let us consider the Hamiltonian:

$$\begin{aligned} &H(x,\ u,\ p,\ q) \\ &\quad := \frac{1}{2}[x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)] + p^T(A(t)x(t) + B(t)u(t)) + q^T A_p(t)x(t). \end{aligned} \tag{2.31}$$

It follows the necessary conditions from the stochastic maximum principle:

$$dp = -\frac{\partial H}{\partial x}dt + qdw = -(Q(t)x(t) + A^T(t)p + A_p^T(t)q)dt + qdw(t). \tag{2.32}$$

with boundary condition:

$$p(T_f) = Q(T_f)x(T_f).$$

The optimal control input:

$$\frac{\partial H}{\partial u} = R(t)u^*(t) + B^T(t)p = 0, \text{ or, } u^*(t) = -R^{-1}(t)B^T(t)p. \tag{2.33}$$

Target: To find p from (2.32) which is still a stochastic differential equation, cannot solve it directly. Ito's lemma is introduced to solve this problem.

Assume that $p(t)$ and $x(t)$ are related by $p(t) = \theta(t, x(t))$. Now using Ito's lemma to $\theta(t, x(t))$ for the given stochastic differential equation (2.19a), we have

$$d\theta = [\theta_t + \theta_x(A(t)x(t) - B(t)R^{-1}(t)B^T(t)\theta) + \frac{1}{2}x^T A_p^T(t)\theta_{xx}A_p(t)x(t)]dt + \theta_x A_p(t)x(t)dw(t). \tag{2.34}$$

Comparing equation (2.32) and (2.34) we get:

$$\begin{aligned} \theta_t + \theta_x(A(t)x(t) - B(t)R^{-1}(t)B^T(t)\theta) + \frac{1}{2}x^T(t)A_p^T(t)\theta_{xx}A_p(t)x(t) \\ = -(Q(t)x(t) + A^T(t)\theta + A_p^T(t)q), \end{aligned} \tag{2.35a}$$

$$q = \theta_x A_p(t)x(t) \tag{2.35b}$$

with boundary condition:

$$\theta(T_f,\ x(T_f)) = Q(T_f)x(T_f).$$

Then, the equation (2.35a) and (2.35b) can be combined as the following simplified form:

$$\begin{aligned} -\theta_t = &\theta_x(A(t)x(t) - B(t)R^{-1}(t)B^T(t)\theta) + \frac{1}{2}x^T(t)A_p^T(t)\theta_{xx}A_p(t)x(t) \\ &+ (Q(t)x(t) + A^T(t)\theta + A_p^T(t)\theta_x A_p(t)x(t)), \end{aligned} \tag{2.36}$$

with boundary condition:

$$\theta(T_f,\, x(T_f)) = Q(T_f)x(T_f).$$

Now, let us consider $\theta = P(t)x(t)$ with $P^T(t) = P(t) \geq 0$, which implies $\theta_x = P(t)$ and $\theta_{xx} = 0$. On the other hand, $\theta(T_f,\, x(T_f)) = P(T_f)x(T_f) = Q(T_f)x(T_f)$, which implies $P(T_f) = Q(T_f)$. So the equation (2.36) can be transferred as follows:

$$\begin{aligned} -\dot{P}(t)x(t) =& P(t)(A(t)x(t) - B(t)R^{-1}(t)B^T(t)P(t)x(t)) \\ &+ (Q(t)x(t) + A^T(t)P(t)x(t) + A_p^T(t)P(t)A_p(t)x(t)) \end{aligned} \tag{2.37}$$

with boundary condition:

$$P(T_f) = Q(T_f).$$

Canceling $x(t)$ from both sides of (2.37), the following stochastic Riccati differential equation (SRDE)can be obtained:

$$\begin{aligned} -\dot{P}(t) =& A^T(t)P(t) + P(t)A(t) + A_p^T(t)P(t)A_p(t) + Q(t) \\ &- P(t)B(t)R^{-1}(t)B^T(t)P(t), \quad P(T_f) = Q(T_f). \end{aligned} \tag{2.38}$$

Moreover, substituting $v_x = P(t)x(t)$ into equation (2.25), we can obtain the state feedback optimal control input,

$$u^*(t) = -R^{-1}(t)B^T(t)P(t)x(t). \tag{2.39}$$

□

2.2.3 Infinite-horizon case

To derive the result for an infinite horizon case, the following facts are used.

Consider the following stochastic system:

$$dx(t) = [Ax(t) + Bu(t)]\,dt + A_p x(t)dw(t), \quad x(0) = x_0. \tag{2.40}$$

Definition 2.1. *[Chen and Zhang (2004)] The stochastic controlled system* (2.40) *is called stabilizable (in the mean square sense), if there exists a feedback control $u(t) = Kx(t)$, such that for any $x_0 \in \mathbb{R}^n$, the closed-loop system*

$$dx(t) = [A + BK]x(t)dt + A_p x(t)dw(t), \quad x(0) = x_0. \tag{2.41}$$

is asymptotically mean square stable, i.e.,

$$\lim_{t\to\infty} \mathbb{E}[x(t)^T x(t)] = 0, \tag{2.42}$$

where is K a constant matrix

Definition 2.2. *If there exist feedback control*

$$u(t) = Kx(t),$$

such that for any $x(0) = x^0$*, the closed-loop stochastic system* (2.40) *is asymptotically mean-square stable, then the stochastic system is called stabilizable.*

Definition 2.3. *[Chen and Zhang (2004)] Consider the following stochastic system with measurement equation.*

$$dx(t) = Ax(t)dt + A_p x(t)dw(t),\ x(0) = x^0, \tag{2.43a}$$

$$z(t) = Cx(t), \tag{2.43b}$$

where $x(t) \in \mathbb{R}^n$ *is the state vector and* $z(t) \in \mathbb{R}^{n_z}$ *is the output measurement;* $A,\ A_p \in \mathbb{R}^{n\times n}$ *and* $C \in \mathbb{R}^{n_z \times n}$ *are the coefficient matrices. If* $z(t) \equiv 0,\ \forall\, t \geq 0$ *implies* $x^0 = 0$*,* $(A,\ A_p \mid C)$ *is called exactly observable.*

To check the exact observability for the system (2.43) we can find the following obsevability matrix [Zhang and Chen (2004)]:

$$\mathcal{O}_s = \begin{bmatrix} C \\ CA \\ CA_p \\ CA_pA \\ CAA_p \\ CA^2 \\ CA_p^2 \\ \vdots \end{bmatrix}.$$

Then, $(A,\ A_p|C)$ is exactly observable iff $\text{rank}(\mathcal{O}_s) = n$.

Lemma 2.2. *Assume that stochastic system* (2.43a) *is asymptotically mean-square stable. Let us define*

$$J = \mathbb{E}\left[\int_0^\infty x^T(t)C^TCx(t)dt\right]. \tag{2.44}$$

If $(A,\ A_p \mid C)$ *is exactly observable, then* $(A,\ A_p)$ *is stable if the following stochastic algebraic Lyapunov equation:*

$$A^TP + PA + A_p^TPA_p + C^TC = 0, \tag{2.45}$$

has a unique positive definite solution $P = P^T$*. Moreover,*

$$J = E[x^T(0)Px(0)]. \tag{2.46}$$

Proof. Let $V(x(t)) = x(t)^T Px(t)$ be the Lyapunov candidate for the system (2.43a), where P is a symmetric positive semi-definite matrix. Now applying Itô's formula, we obtain

$$\begin{aligned} dV(x(t)) &= V_x Ax(t) + \frac{1}{2} x^T(t) A_p^T V_{xx} A_p x(t) \\ &= x^T(t)[A^T P + PA + A_p^T PA_p]x(t), \end{aligned} \tag{2.47}$$

which is stable if $dV(x(t)) < 0$. If $(A,\ A_p \mid C)$ is exactly observable, we can form the Lyapunov stabilizable equation by integrating and taking expectoration operator $(\mathbb{E}[\cdot])$ in (2.47) as follows:

$$\mathbb{E}\left[\int_0^\infty x^T(t)[A^T P + PA + A_p^T PA_p]x(t)dt\right] = \mathbb{E}\left[\int_0^\infty x^T(t)C^T Cx(t)dt\right], \tag{2.48}$$

i.e.,

$$A^T P + PA + A_p^T PA_p + C^T C = 0. \tag{2.49}$$

Hence, the first part of Lemma 2.2 is proved.

Let us consider the proof of second part.

$$\begin{aligned} &\mathbb{E}\Big[\int_0^t x^T(s)C^T Cx(s)ds\Big] \\ &\quad = -\mathbb{E}\left[\int_0^t x^T(s)[A^T P + PA + A_p^T PA_p]x(s)ds\right] \\ &\quad = -\mathbb{E}\left[\int_0^t x^T(s)\dot{P}x(s)ds\right] \quad \text{[Itô's formula (2.48).]} \\ &\quad = \mathbb{E}[x^T(0)Px(0)] - \mathbb{E}[x^T(t)Px(t)] \\ &\quad = \mathbb{E}[x^T(0)Px(0)], \quad \text{when } t \to \infty. \end{aligned}$$

Hence the Lemma 2.2 is proved. □

Now, consider the next stochastic linear quadratic optimal control problem in the case of infinite-horizon:

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.50a}$$

$$J(x_0,\ u) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \big(x^T(t)Qx(t) + u^T(t)Ru(t)\big)\, dt\right], \tag{2.50b}$$

where $A, B, A_p, Q = Q^T \geq 0, R = R^T > 0$ are the coefficient matrices of suitable dimensions. In order to solve the above-mentioned optimal control problem, the following result can be obtained.

Theorem 2.3. *For the stochastic optimal control problem* (2.50)*, suppose that the following stochastic algebraic Riccati equation (SARE) has the solution* $P^T = P \geq 0$*:*

$$PA + A^T P + Q - PBR^{-1}B^T P + A_p^T PA_p = 0, \tag{2.51}$$

then the optimal control problem admits a state feedback solution,

$$u^*(t) = Kx(t) = -R^{-1}B^T Px(t). \tag{2.52}$$

Proof. Recall SRDE (2.20) of Theorem 2.2. As $T_f \to \infty$, $P(t)$ approaches to *steady-state*. Therefore, $\dot{P}(t) = 0$. Then form equation (2.20), we can obtain the following stochastic algebraic Riccati equation (SARE):

$$PA + A^T P + Q - PBR^{-1}B^T P + A_p^T PA_p = 0. \tag{2.53}$$

Moreover, using the same technique as finite-horizon case, the optimal state feedback control can be derived as follows:

$$u^*(t) = Kx(t) = -R^{-1}B^T Px(t). \tag{2.54}$$

For Lyapunov stability analysis, consider SARE (2.53) as the form of stochastic algebraic Lyapunov equation:

$$(A+BK)^T P + P(A+BK) + (Q + K^T RK) + A_p^T PA_p = 0, \tag{2.55}$$

with $K = -R^{-1}B^T P$. By Lemma 2.2, if $(A+BK,\ A_p \mid \sqrt{Q+K^T RK})$ is exactly observable, then $(A+BK,\ A_p)$ is stable if the stochastic algebraic Lyapunov equation (2.55) has a unique positive definite solution $P = P^T$. □

2.2.4 Numerical examples

Finite-horizon case

Consider the linear stochastic differential equation:

$$dx(t) = (-x(t) + u(t))dt + \varepsilon x(t)dw(t), \quad x(0) = 1, \tag{2.56a}$$

$$J(x_0,\ u) := \mathbb{E}\left[\int_0^1 (x^2(t) + u^2(t))dt\right]. \tag{2.56b}$$

To find the optimal control for the above problem. Hamiltonian is defined by

$$H = x^2 + u^2 + p(-x+u) + q\varepsilon x. \tag{2.57}$$

It follows the necessary conditions from the stochastic maximum principle:

$$dx^*(t) = \frac{\partial H}{\partial p}dt + \varepsilon x^*(t)dw(t) = (-x^*(t) + u(t))dt + \varepsilon x^*(t)dw(t). \tag{2.58}$$

with initial condition $x^*(0) = 1$.

$$dp = -\frac{\partial H}{\partial x(t)}dt + qdw(t) = -(2x(t) - p + q\varepsilon)dt + qdw(t). \tag{2.59}$$

with boundary condition: $p(1) = 0$.
The optimal control input:

$$\frac{\partial H}{\partial u(t)} = 2u^*(t) + p = 0, \text{or, } u^*(t) = -\frac{p}{2}. \tag{2.60}$$

To find p from (2.59) which is still a stochastic differential equation, cannot solve it directly.

Assume that $p(t)$ and $x^*(t)$ are related by $p(t) = \theta(t, x^*(t))$. Now using Ito's lemma to $\theta(t, x^*(t))$, we have (omitting * from x^* for the simplicity of notation)

$$d\theta = [\theta_t + \theta_x\left(-x(t) - \frac{\theta}{2}\right) + \frac{1}{2}\theta_{xx}(\varepsilon x(t))^2]dt + \theta_x \varepsilon x(t)dx. \tag{2.61}$$

Comparing equation (2.59) and (2.61) we get:

$$\theta_t + \theta_x\left(-x(t) - \frac{\theta}{2}\right) + \frac{1}{2}\theta_{xx}(\varepsilon x(t))^2 = -(2x(t) - \theta + q\varepsilon), \tag{2.62a}$$

$$q = \theta_x \varepsilon x(t) \tag{2.62b}$$

with boundary condition: $\theta(1, x) = 0$. Then, the equation (2.62) means the following simplified form:

$$-\theta_t = 2x(t) - \theta + \theta_x \varepsilon^2 x(t) + \theta_x\left(-x(t) - \frac{\theta}{2}\right) + \frac{1}{2}\theta_{xx}(\varepsilon x(t))^2. \tag{2.63}$$

with boundary condition: $\theta(1, x) = 0$. Since (2.63) is a deterministic partial differential equation, so we can solve it numerically by using backward difference formula. This provides the solution for the control input in an open-loop pattern.

For Closed-loop pattern, let $\theta = zx$, or, $\theta_x = z$ and $\theta_{xx} = 0$.
On the other hand, $\theta(1, x) = z(1)x = 2x$or, $z(1) = 2$.
So the equation (2.63) takes the for

$$-\dot{z}x = 2x - zx + z\varepsilon^2 x - zx - \frac{z^2 x}{2}$$

$$\text{or, } \dot{z} = \frac{z^2}{2} + (2-\varepsilon^2)z - 2; \quad z(1) = 2. \tag{2.64}$$

We can solve the differential equation (2.64) numerically by backward difference formula.

Solving the system (2.64) for z, we can find the state feedback optimal control from (2.60),

$$u^*(t) = -p/2 = -(z/2)x(t). \tag{2.65}$$

The state trajectory can be depicted by Fig. 2.1 and can be detected by the following equation:

$$x(t+h) = x(t) + h(-1-z/2)x(t) + \varepsilon x(t)\sqrt{h}N(0,1), \quad x(0) = 1. \tag{2.66}$$

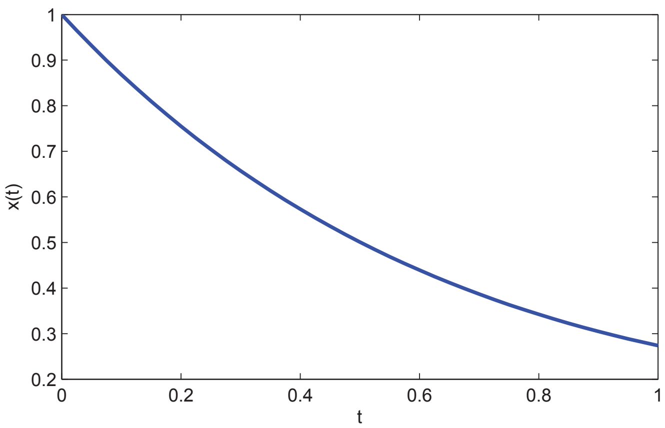


Fig. 2.1: Finite-horizon state trajectory.

Infinite-horizon case

Recall the stochastic linear quadratic optimal control problem (2.50). Let us consider the following system matrices:

$$A = \begin{bmatrix} -2.98 & 0.93 & 0 & -0.034 \\ -0.99 & -0.21 & 0.035 & -0.0011 \\ 0 & 0 & 0 & 1 \\ 0.39 & -5.555 & 0 & -1.89 \end{bmatrix}, \ B = \begin{bmatrix} -0.032 \\ 0 \\ 0 \\ -1.6 \end{bmatrix}, \ x_0 = \begin{bmatrix} 1 \\ 0.5 \\ -0.5 \\ -1 \end{bmatrix},$$
$$A_p = 0.1A, \ Q = \textbf{diag}\begin{bmatrix} 1.5 & 0.8 & 2.3 & 1.9 \end{bmatrix}, \ R = \textbf{diag}\begin{bmatrix} 2.5 & 1.8 & 1.3 & 0.9 \end{bmatrix}.$$

In order to solve this optimal control problem let us consider the SARE (2.51) as the following nonlinear matrix function:

$$\mathscr{G}(P) = PA + A^T P + Q - PBR^{-1}B^T P + A_p^T P A_p = 0. \tag{2.67}$$

To solve the nonlinear matrix function (2.67) Newton's method can be applied as follows:

Newton's Method

The iterative form of the SARE (2.67) by Newton's iteration is:

$$P^{(k+1)}(A-SP^{(k)})+(A-SP^{(k)})^T P^{(k+1)}+A_p^T P^{(k+1)}A_p+P^{(k)}SP^{(k)}+Q=0, \qquad (2.68)$$

where $S=BR^{-1}B^T$ and $k=0,\ 1,\ \ldots$.

This result can be established using Newton's method as follows. By the definition of Newton's method, the following equation holds:

$$\begin{aligned}\mathrm{vec}P^{(k+1)} =&\mathrm{vec}P^{(k)}-\left[\frac{\partial \mathrm{vec}\mathscr{G}(P)}{\partial(\mathrm{vec}P)^T}\bigg|_{P=P^{(k)}}\right]^{-1}\mathrm{vec}\mathscr{G}(P^{(k)})\\ &=\mathrm{vec}P^{(k)}-\left[(A-SP^{(k)})^T\otimes I_n+I_n\otimes(A-SP^{(k)})^T+A_p^T\otimes A_p^T\right]^{-1}\mathrm{vec}\mathscr{G}(P^{(k)}).\end{aligned} \qquad (2.69)$$

Thus, the operation $(A-SP^{(k)})^T\otimes I_n+I_n\otimes(A-SP^{(k)})^T+A_p^T\otimes A_p^T$ yields

$$\left[(A-SP^{(k)})^T\otimes I_n+I_n\otimes(A-SP^{(k)})^T+A_p^T\otimes A_p^T\right] \qquad (2.70)$$

$$\times(\mathrm{vec}P^{(k+1)}-\mathrm{vec}P^{(k)})+\mathrm{vec}\mathscr{G}(P^{(k)})=0. \qquad (2.71)$$

Moreover, using the formulation $\mathrm{vec}(AXB)=[B^T\otimes A]\mathrm{vec}X$ in the left hand side of (2.71), we obtain

$$\begin{aligned}\mathrm{LHS}=&\left[(A-SP^{(k)})^T\otimes I_n+I_n\otimes(A-SP^{(k)})^T+A_p^T\otimes A_p^T\right]\\ &\times(\mathrm{vec}P^{(k+1)}-\mathrm{vec}P^{(k)})+\mathrm{vec}\mathscr{G}(P^{(k)})\\ =&\,\mathrm{vec}\Big[(P^{(k+1)}-P^{(k)})(A-SP^{(k)})+(A-SP^{(k)})^T(P^{(k+1)}-P^{(k)})\\ &+A_p^T(P^{(k+1)}-P^{(k)})A_p\Big]+\mathrm{vec}\mathscr{G}(P^{(k)})\\ =&\,\mathrm{vec}\left[P^{(k+1)}(A-SP^{(k)})+(A-SP^{(k)})^TP^{(k+1)}+A_p^TP^{(k+1)}A_p\right]\\ &-\mathrm{vec}\left[P^{(k)}(A-SP^{(k)})+(A-SP^{(k)})^TP^{(k)}+A_p^TP^{(k)}A_p\right]+\mathrm{vec}\mathscr{G}(P^{(k)})\\ =&\,\mathrm{vec}\left[P^{(k+1)}(A-SP^{(k)})+(A-SP^{(k)})^TP^{(k+1)}+A_p^TP^{(k+1)}A_p+P^{(k)}SP^{(k)}+Q\right]\\ =&\,0\ (\mathrm{RHS}).\end{aligned}$$

which is the desired result.

Theorem 2.4. *Newton-Kantorovich theorem [Yamamoto 1986, Ortega 1990] : Assume that*

$$G : \mathbf{R}^n \rightarrow \mathbf{R}^n \tag{2.72}$$

is differentiable on a convex set D. Suppose that the inverse of map G exists and moreover it is differentiable on set D and that

$$||G'(\mathbf{x}) - G'(\mathbf{y})|| \leq \gamma ||\mathbf{x} - \mathbf{y}|| \tag{2.73}$$

for all $\mathbf{x},\ \mathbf{y} \in D$*.*

Proof. Suppose that there is an $\mathbf{x}^0 \in D$ such that

$$||G'(\mathbf{x}^0)^{-1}|| \leq \beta \tag{2.74}$$

$$||G'(\mathbf{x}^0)^{-1} G(\mathbf{x}^0)|| \leq \eta \tag{2.75}$$

and

$$\theta := \beta\gamma\eta < 1/2 \tag{2.76}$$

Assume that

$$S := \{\ \mathbf{x} :\ \ ||\mathbf{x} - \mathbf{x}^0|| \leq t^*\ \} \subset D \tag{2.77}$$

and

$$t^* = \frac{1 - \sqrt{1 - 2\theta}}{\beta\gamma} \tag{2.78}$$

Then Newton iterations

$$\mathbf{x}^{k+1} = \mathbf{x}^k - G'(\mathbf{x}^k)^{-1} G(\mathbf{x}^k), \tag{2.79}$$

$k = 0,\ 1,\ \cdots$ are well defined and converge to a solution $\mathbf{x}^*$ of $G(\mathbf{x}) = 0$ in S. Moreover, the solution $\mathbf{x}^*$ is unique in $\tilde{S} \cap D$, where

$$\tilde{S} := \{\ \mathbf{x} :\ \ ||\mathbf{x} - \mathbf{x}^0|| \leq \tilde{t}\ \} \subset D \tag{2.80}$$

$$\tilde{t} = \frac{1 + \sqrt{1 - 2\theta}}{\beta\gamma} \tag{2.81}$$

and error estimate is given by

$$||\mathbf{x}^* - \mathbf{x}^k|| \leq \frac{(2\theta)^{2^k}}{2^k \beta\gamma} = 2^{1-k}(2\theta)^{2^k - 1}\eta,\ k = 0,\ 1,\ ... \tag{2.82}$$

Hence, Newton-Kantorovich theorem is proved. □

Newton's algorithm

Inputs: Let $P = P^{(0)}$ be the given initial matrix; $ITER$ is the maximum number of iterations; TOL is the tolerance of convergence.
Output: Solution matrix P.

Step 1 For $k = 1,\ 2, \cdots,\ ITER$ do **Step 2** to **Step 3**.

Step 2 Calculate the following newtons formula:

$$\text{vec}P^{(k+1)} = \text{vec}P^{(k)} - \left[\frac{\partial \text{vec}\mathscr{G}(P)}{\partial (\text{vec}P)^T} \bigg|_{P=P^{(k)}} \right]^{-1} \text{vec}\mathscr{G}(P^{(k)}), \tag{2.83}$$

where

$$\frac{\partial \text{vec}\mathscr{G}(P)}{\partial (\text{vec}P)^T} = A^T \otimes I_n + I_n \otimes A^T - I_n \otimes PBR^{-1}B^T - BR^{-1}B^T P \otimes I_n + A_p^T \otimes A_p^T.$$

Step 3 If $\|P^{(k+1)} - P^{(k)}\| < TOL$, **stop**.

Step 4: Output

Step 5: End

This task can also be accomplished by using the Lyapunov iterative technique in **Step 2** of the above algorithm. The MATLAB built-in command `lyap` is very useful for this kind of simulation. The application of this algorithm through MATLAB simulation provides the following results:

$$P = \begin{bmatrix} 2.7357 & -7.5039 & 1.3446 & 0.7675 \\ -7.5039 & 23.8648 & -4.9545 & -2.9577 \\ 1.3446 & -4.9545 & 4.0786 & 1.3543 \\ 0.7675 & -2.9577 & 1.3543 & 0.9572 \end{bmatrix},$$

$$K = \begin{bmatrix} 0.5262 & -1.9890 & 0.8840 & 0.6224 \end{bmatrix}.$$

k	Lyapunov method	Newton's method
0	2.4970	2.4970
1	3.0127×10^{-1}	3.1088×10^{-1}
2	1.3197×10^{-3}	1.0294×10^{-3}
3	2.2233×10^{-6}	4.1651×10^{-9}
4	5.886×10^{-9}	6.4285×10^{-15}
5	1.7001×10^{-11}	
6	1.2671×10^{-13}	
7	6.7299×10^{-14}	

Table 2.1: Error in each iteration.

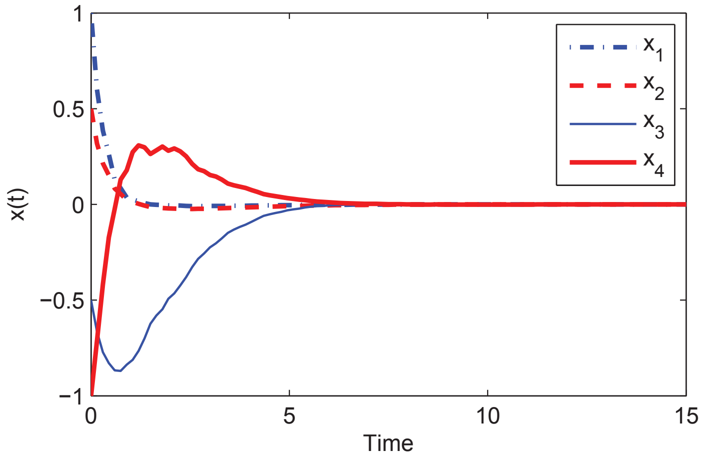


Fig. 2.2: Trajectory of the state.

From Table 2.1, it can be observed that the algorithm converges to the exact solution with an accuracy of $\|\mathscr{G}(P^{(k)})\| < 10^{-13}$ after seven iterations using the Lyapunov iterative technique. It can also be observed that Newton's method attains quadratic convergence only after four iterations under the appropriate initial conditions. Therefore, Newton's method is potentially fast and more accurate than the widely used the Lyapunov iterative technique.

Fig. 2.2 shows the response of the system with a state trajectory. It shows that the state variables $x(k)$ can stabilize the given system, which implies that the proposed method is very useful and reliable.

2.3 Solution concepts for multi-player problems

To understand the multi-player situation, let us consider the two-player game problems for the cooperative and non-cooperative cases. Stochastic Pareto optimality and Nash equilibrium solution concepts are introduced for cooperative and non-cooperative game problems, respectively. Let us consider the linear stochastic system of two players:

$$dx(t) = [Ax(t) + B_1u_1(t) + B_2u_2(t)]dt + A_px(t)dw(t), \quad x(0) = x^0, \tag{2.84}$$

and the cost functionals are

$$J_1(x_0,\ u_1,\ u_2) = \frac{1}{2}\mathbb{E}\left[\int_0^{T_f} \left(x^T(t)Q_1x(t) + u_1^T(t)R_{11}u_1(t) + u_2^T(t)R_{12}u_2(t)\right) dt\right], \tag{2.85a}$$

$$J_2(x_0,\ u_1,\ u_2) = \frac{1}{2}\mathbb{E}\left[\int_0^{T_f} \left(x^T(t)Q_2x(t) + u_1^T(t)R_{21}u_1(t) + u_2^T(t)R_{22}u_2(t)\right) dt\right], \tag{2.85b}$$

where $Q_i = Q_i^T \geq 0$, $R_{ij} = R_{ij}^T \geq 0$ for $i \neq j$ and $R_{ii} = R_{ii}^T > 0$, $i, j = 1,\ 2$.

2.3.1 Pareto optimal solution

Let us consider the following linear stochastic system with linear quadratic cost functions,

$$dx(t) = [Ax(t) + \sum_{i=1}^{N} B_i u_i(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.86a}$$

$$J_i(x_0,\ u_1,\ \dots\ ,u_N) := \frac{1}{2}\mathbb{E}\left[\int_0^{T_f}\left[x^T(t)Q_i x(t) + \sum_{j=1}^{N} u_j^T(t)R_{ij}u_j(t)\right]dt\right], \tag{2.86b}$$

where $Q_i = Q_i^T \geq 0$, $R_{ij} = R_{ij}^T \geq 0$ for $i \neq j$ and $R_{ii} = R_{ii}^T > 0$, $i, j = 1,\ \dots\ ,N$.

Definition 2.4. *A strategy-set $(u_1,\ \cdots\ ,u_N)$ is said to be a Pareto optimal strategy if it minimizes a sum of the cost of functional of all players denoted by*

$$J(u_1,\ \cdots\ ,u_N) = \sum_{i=1}^{N} r_i J_i(x_0,\ u_1,\ \cdots\ ,u_N), \tag{2.87}$$

where $\sum_{i=1}^{N} r_i = 1$ for some $0 < r_i < 1$.

Theorem 2.5. *For the stochastic optimal control problem* (2.86)*, suppose that the following stochastic Riccati differential equation (SRDE) has the solution $P^T(t) = P(t) \geq 0$:*

$$-\dot{P}(t) = P(t)A + A^T P(t) + \mathbf{Q} - P(t)\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P(t) + A_p^T P(t)A_p, \quad P(T_f) = 0, \tag{2.88}$$

then the Pareto optimal control problem admits a state feedback solution,

$$u^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)P(t)x(t), \tag{2.89}$$

where

$$\begin{aligned}
&\mathbf{B} := \left[B_1,\ \dots\ ,B_N\right],\\
&u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{bmatrix},\\
&\mathbf{Q} = \sum_{i=1}^{N} r_i Q_i,\\
&\mathbf{R} = \textbf{block diag}\left[\textstyle\sum_{i=1}^{N} r_i R_{i1} \quad \dots \quad \sum_{i=1}^{N} r_i R_{iN}\right].
\end{aligned}$$

Proof. If we centralized the system (2.86) base on the Definition 2.4, we can can rewrite it as follows:

$$dx(t) = [Ax(t) + \mathbf{B}u(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.90a}$$

$$J(x_0,\ u) := \frac{1}{2}\mathbb{E}\left[\int_0^{T_f} \left(x^T(t)\mathbf{Q}x(t) + u^T(t)\mathbf{R}u(t)\right) dt\right], \tag{2.90b}$$

where

$$\begin{aligned}
&\mathbf{B} := \left[B_1,\ \dots\ ,B_N\right], \\
&u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{bmatrix}, \\
&\mathbf{Q} = \sum_{i=1}^{N} r_i Q_i, \\
&\mathbf{R} = \textbf{block diag}\left[\sum_{i=1}^N r_i R_{i1} \quad \dots \quad \sum_{i=1}^N r_i R_{iN}\right].
\end{aligned}$$

We define the finite-horizon value function,

$$v(t,x) = \frac{1}{2} \min_{u(t)\in U} \mathbb{E}\left[\int_t^{T_f} \{x^T(s)\mathbf{Q}x(s) + u^T(s)\mathbf{R}u(s)\} ds\right]. \tag{2.91}$$

It will satisfy the stochastic Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned}
-v_t = \min_{u(t)\in U} [&\frac{1}{2}(x^T(t)\mathbf{Q}x(t) + u^T(t)\mathbf{R}u(t)) + v_x^T(A(t)x(t) + \mathbf{B}u(t)) \\
&+ \frac{1}{2}(A_p(t)x(t))^T(A_p(t)x(t))v_{xx}],
\end{aligned} \tag{2.92}$$

with boundary condition: $v(T_f,x) = 0$.

To minimize the right hand side of (2.92) with respect to the control input $u(t)$ we get,

$$\mathbf{R}u^*(t) + \mathbf{B}^T v_x = 0, \tag{2.93}$$

or,

$$u^*(t) = -\mathbf{R}^{-1}\mathbf{B}^T v_x. \tag{2.94}$$

Now, if we insert this optimal state feedback control input $u^*(t)$ into equation (2.92) we get,

$$\begin{aligned}
-v_t =& \frac{1}{2}(x^T(t)\mathbf{Q}x(t) + v_x^T\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T v_x) + v_x^T(A(t)x(t) - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T v_x) \\
&+ \frac{1}{2}x^T(t)A_p^T(t)A_p(t)x(t)v_{xx}, \\
=& \frac{1}{2}(x^T(t)\mathbf{Q}x(t) - v_x^T\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T v_x) + v_x^T A(t)x(t) + \frac{1}{2}x^T(t)A_p^T(t)A_p(t)x(t)v_{xx},
\end{aligned} \tag{2.95}$$

with boundary condition:

$$v(T_f,x) = 0.$$

Now let

$$v(t,x) = \frac{1}{2}x^T(t)P(t)x(t), \tag{2.96}$$

where $P(t)$ is a symmetric positive semidefinite matrix. Therefore,

$$v_x = P(t)x(t),$$

and

$$v_{xx} = P(t).$$

So, from (2.95) we get,

$$\begin{aligned}
-\frac{1}{2}x^T(t)\dot{P}(t)x(t) =& \frac{1}{2}(x^T(t)\mathbf{Q}x(t) - x^T(t)P(t)\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P(t)x(t)) \\
&+ \underbrace{x^T(t)P(t)A(t)x(t)}_{\frac{1}{2}x^T(t)(P(t)A(t)+A^T(t)P(t))x(t)} + \frac{1}{2}x^T(t)A_p^T(t)P(t)A_p(t)x(t),
\end{aligned} \tag{2.97}$$

with boundary condition: $P(T_f) = 0$. Simply,

$$-\dot{P}(t) = P(t)A(t) + A^T(t)P(t) + \mathbf{Q} - P(t)\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P(t) + A_p^T(t)P(t)A_p(t), \quad P(T_f) = 0. \tag{2.98}$$

Moreover, substituting $v_x = P(t)x(t)$ into equation (2.94), we can obtain the state feedback optimal control input,

$$u^*(t) = -\mathbf{R}^{-1}\mathbf{B}^T P(t)x(t). \tag{2.99}$$

□

Infinite-horizon case:

Let us consider the following linear stochastic system with linear quadratic cost functions,

$$dx(t) = [Ax(t) + \sum_{i=1}^{N} B_i u_i(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.100a}$$

$$J_i(x_0,\, u_1,\, \dots\, ,u_N) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left[x^T(t)Q_i x(t) + \sum_{j=1}^{N} u_j^T(t)R_{ij}u_j(t)\right]dt\right], \tag{2.100b}$$

where $Q_i = Q_i^T \geq 0$, $R_{ij} = R_{ij}^T \geq 0$ for $i \neq j$ and $R_{ii} = R_{ii}^T > 0$, $i, j = 1,\, \dots\, ,N$.

Theorem 2.6. *For the stochastic optimal control problem* (2.100)*, suppose that the following stochastic algebraic Riccati equation (SARE) has the solution* $P^T = P \geq 0$*:*

$$PA + A^T P + \mathbf{Q} - P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P + A_p^T P A_p = 0, \tag{2.101}$$

then the optimal control problem admits a state feedback solution,

$$u^*(t) = Kx(t) = -\mathbf{R}^{-1}\mathbf{B}^T Px(t). \tag{2.102}$$

Proof. Recall SRDE (2.88) of Theorem 2.5. As $T_f \to \infty$, $P(t)$ approaches to *steady-state*. Therefore, $\dot{P}(t) = 0$. Then form equation (2.88), we can obtain the following stochastic algebraic Riccati equation (SARE):

$$PA + A^T P + \mathbf{Q} - P\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T P + A_p^T P A_p = 0. \tag{2.103}$$

Moreover, using the same technique as finite-horizon case, the optimal state feedback control can be derived as follows:

$$u^*(t) = Kx(t) = -\mathbf{R}^{-1}\mathbf{B}^T P x(t). \tag{2.104}$$

For Lyapunov stability analysis, consider SARE (2.103) as the form of stochastic algebraic Lyapunov equation:

$$(A + \mathbf{B}K)^T P + P(A + \mathbf{B}K) + (\mathbf{Q} + K^T \mathbf{R} K) + A_p^T P A_p = 0, \tag{2.105}$$

with $K = -\mathbf{R}^{-1}\mathbf{B}^T P$. By Lemma 2.2, if $(A + \mathbf{B}K,\ A_p \mid \sqrt{\mathbf{Q} + K^T \mathbf{R} K})$ is exactly observable, then $(A + \mathbf{B}K,\ A_p)$ is stable if the stochastic algebraic Lyapunov equation (2.105) has a unique positive definite solution $P = P^T$. □

2.3.2 Nash equilibrium solution

Recall the two-player game problem (2.84)–(2.85). We call the pair $(u_1,\ u_2) \in U_1 \times U_2$ a Nash equilibrium if

$$J_1(x_0,\ u_1^*,\ u_2^*) \le J_1(x_0,\ u_1,\ u_2^*), \quad \forall u_1 \in U_1, \tag{2.106a}$$

$$J_2(x_0,\ u_1^*,\ u_2^*) \le J_2(x_0,\ u_1^*,\ u_2), \quad \forall u_2 \in U_2. \tag{2.106b}$$

Theorem 2.7. *If there exist two symmetric positive semi-definite matrices $P_1(t)$ and $P_2(t)$ satisfying the following cross-coupling stochastic Riccati differential equations (SRDEs):*

$$\begin{aligned} -\dot{P}_1(t) =& (A + B_2K_2)^T P_1(t) + P_1(t)(A + B_2K_2) + A_p^T P_1 A_p + Q_1 + K_2^T R_{12} K_2 \\ & - P_1(t) B_1 R_{11}^{-1} B_1^T P_1(t), \quad P_1(T_f) = 0, \end{aligned} \tag{2.107a}$$

$$\begin{aligned} -\dot{P}_2(t) =& (A + B_1K_1)^T P_2(t) + P_2(t)(A + B_1K_1) + A_p^T P_2 A_p + Q_2 + K_1^T R_{21} K_1 \\ & - P_2(t) B_2 R_{22}^{-1} B_2^T P_2(t), \quad P_2(T_f) = 0, \end{aligned} \tag{2.107b}$$

then, the state feedback strategy pair $(u_1^*(t),\ u_2^*(t))$ is a Nash equilibrium for the system (2.84)–(2.85) *, where*

$$u_1^*(t) = K_1 x(t) = -R_{11}^{-1} B_1^T P_1(t) x(t), \tag{2.108a}$$

$$u_2^*(t) = K_2 x(t) = -R_{22}^{-1} B_2^T P_2(t) x(t). \tag{2.108b}$$

Proof. Substituting $u_2^*(t) = K_2 x(t)$ into state equation (2.84) and the cost functional for the first player (2.85a) gives

$$dx(t) = [(A + B_2K_2)x(t) + B_1u_1(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.109a}$$

$$J_1(x_0,\, u_1,\, u_2^*) = \frac{1}{2}\mathbb{E}\Big[\int_0^{T_f} \big(x^T(t)Q_1x(t) + u_1^T(t)R_{11}u_1(t) + x^T(t)K_2^T R_{12}K_2x(t)\big)dt\Big]. \tag{2.109b}$$

Hence, by Theorem 2.2, if the following stochastic Riccati differential equation (SRDE) has the solution $P_1^T(t) = P_1(t) \geq 0$:

$$\begin{aligned} -\dot{P}_1(t) =& (A + B_2K_2)^T P_1(t) + P_1(t)(A + B_2K_2) + A_p^T P_1 A_p + Q_1 + K_2^T R_{12}K_2 \\ & - P_1(t)B_1R_{11}^{-1}B_1^T P_1(t), \quad P_1(T_f) = 0, \end{aligned} \tag{2.110}$$

then the stochastic optimal control problem (2.109) admits a state feedback solution,

$$u_1^*(t) = K_1x(t) = -R_{11}^{-1}B_1^T P_1(t)x(t). \tag{2.111}$$

Conversely, substituting $u_1^*(t) = K_1x(t)$ into state equation (2.84) and the cost functional for the first player (2.85b) gives

$$dx(t) = [(A + B_1K_1)x(t) + B_2u_2(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.112a}$$

$$J_2(x_0,\, u_1^*,\, u_2) = \frac{1}{2}\mathbb{E}\Big[\int_0^{T_f} \big(x^T(t)Q_2x(t) + u_2^T(t)R_{22}u_2(t) + x^T(t)K_1^T R_{21}K_1x(t)\big)dt\Big]. \tag{2.112b}$$

Hence, by Theorem 2.2, if the following stochastic Riccati differential equation (SRDE) has the solution $P_2^T(t) = P_2(t) \geq 0$:

$$\begin{aligned} -\dot{P}_2(t) =& (A + B_1K_1)^T P_2(t) + P_2(t)(A + B_1K_1) + A_p^T P_2 A_p + Q_2 + K_1^T R_{21}K_1 \\ & - P_2(t)B_2R_{22}^{-1}B_2^T P_2(t), \quad P_2(T_f) = 0, \end{aligned} \tag{2.113}$$

then the stochastic optimal control problem (2.112) admits a state feedback solution,

$$u_2^*(t) = K_2x(t) = -R_{22}^{-1}B_2 T P_2(t)x(t). \tag{2.114}$$

Hence the theorem is proved. □

Infinite horizon case

In the case of infinite horizon, let us consider the linear stochastic system of two players:

$$dx(t) = [Ax(t) + B_1u_1(t) + B_2u_2(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{2.115}$$

and the cost functionals are

$$J_1(x_0,\ u_1,\ u_2) = \frac{1}{2}\mathbb{E}\left[\int_0^{\infty} \left(x^T(t)Q_1x(t) + u_1^T(t)R_{11}u_1(t) + u_2^T(t)R_{12}u_2(t)\right) dt\right], \quad (2.116a)$$

$$J_2(x_0,\ u_1,\ u_2) = \frac{1}{2}\mathbb{E}\left[\int_0^{\infty} \left(x^T(t)Q_2x(t) + u_2^T(t)R_{22}u_2(t) + u_1^T(t)R_{21}u_1(t)\right) dt\right], \quad (2.116b)$$

where $Q_i = Q_i^T \geq 0$, $R_{ij} = R_{ij}^T \geq 0$ for $i \neq j$ and $R_{ii} = R_{ii}^T > 0$, $i, j = 1,\ 2$.

Lemma 2.3. *If there exist two symmetric positive semi-definite matrices P_1 and P_2 satisfying the following cross-coupling stochastic Riccati differential equations (SRDEs):*

$$(A + B_2K_2)^T P_1 + P_1(A + B_2K_2) + A_p^T P_1 A_p + Q_1 + K_2^T R_{12} K_2 - P_1 B_1 R_{11}^{-1} B_1^T P_1 = 0, \quad (2.117a)$$

$$(A + B_1K_1)^T P_2 + P_2(A + B_1K_1) + A_p^T P_2 A_p + Q_2 + K_1^T R_{21} K_1 - P_2 B_2 R_{22}^{-1} B_2^T P_2 = 0, \quad (2.117b)$$

then, the state feedback strategy pair $(u_1^*(t),\ u_2^*(t))$ is a Nash equilibrium for the system (2.115)–(2.116)*, where*

$$u_1^*(t) = K_1x(t) = -R_{11}^{-1}B_1^T P_1 x(t), \quad (2.118a)$$

$$u_2^*(t) = K_2x(t) = -R_{22}^{-1}B_2^T P_2 x(t). \quad (2.118b)$$

Proof. Substituting $u_2^*(t) = K_2x(t)$ into state equation (2.115) and the cost functional for the first player (2.116a) gives

$$dx(t) = [(A + B_2K_2)x(t) + B_1u_1(t)]dt + A_px(t)dw(t), \quad x(0) = x^0, \quad (2.119a)$$

$$J_1(x_0,\ u_1,\ u_2^*) = \frac{1}{2}\mathbb{E}\Big[\int_0^{\infty} \left(x^T(t)Q_1x(t) + u_1^T(t)R_{11}u_1(t) + x^T(t)K_2^T R_{12}K_2x(t)\right)dt\Big]. \quad (2.119b)$$

Hence, by Theorem 2.3, if the following stochastic Riccati differential equation (SRDE) has the solution $P_1^T = P_1 \geq 0$:

$$(A + B_2K_2)^T P_1 + P_1(A + B_2K_2) + A_p^T P_1 A_p + Q_1 + K_2^T R_{12} K_2 - P_1 B_1 R_{11}^{-1} B_1^T P_1 = 0, \quad (2.120)$$

then the stochastic optimal control problem (2.119) admits a state feedback solution,

$$u_1^*(t) = K_1x(t) = -R_{11}^{-1}B_1^T P_1 x(t). \quad (2.121)$$

Conversely, substituting $u_1^*(t) = K_1x(t)$ into state equation (2.115) and the cost functional for the first player (2.116b) gives

$$dx(t) = [(A + B_1K_1)x(t) + B_2u_2(t)]dt + A_px(t)dw(t), \quad x(0) = x^0, \quad (2.122a)$$

$$J_2(x_0,\ u_1^*,\ u_2) = \frac{1}{2}\mathbb{E}\Big[\int_0^\infty \big(x^T(t)Q_2x(t) + u_2^T(t)R_{22}u_2(t) + x^T(t)K_1^TR_{21}K_1x(t)\big)dt\Big]. \tag{2.122b}$$

Hence, by Theorem 2.3, if the following stochastic Riccati differential equation (SRDE) has the solution $P_2^T = P_2 \geq 0$:

$$(A+B_1K_1)^TP_2 + P_2(A+B_1K_1) + A_p^TP_2A_p + Q_2 + K_1^TR_{21}K_1 - P_2B_2R_{22}^{-1}B_2^TP_2 = 0, \tag{2.123}$$

then the stochastic optimal control problem (2.122) admits a state feedback solution,

$$u_2^*(t) = K_2x(t) = -R_{22}^{-1}B_2TP_2x(t). \tag{2.124}$$

Hence the lemma is proved. □

2.3.3 Numerical examples

Finite-horizon case:

Consider a linear stochastic two-player Nash equilibrium problem:

$$dx(t) = (-x+u_1+u_2)dt + \varepsilon x dw(t), \quad x(0) = 1, \tag{2.125}$$

with cost functionals

$$J_1 = \frac{1}{2}\mathbb{E}\left[\int_0^1 (x^2+u_1^2+2u_2^2)dt\right], \tag{2.126a}$$

$$J_2 = \frac{1}{2}\mathbb{E}\left[\int_0^1 (x^2+2u_1^2+u_2^2)dt\right]. \tag{2.126b}$$

To find the optimal control for the above problem.

We call a pair $(u_1,\ u_2) \in U_1 \times U_2$ a Nash equilibrium if

$$J_1(u_1^*,\ u_2^*) \leq J_1(u_1,\ u_2^*), \quad \forall u_1 \in U_1, \tag{2.127a}$$

$$J_2(u_1^*,\ u_2^*) \leq J_2(u_1^*,\ u_2), \quad \forall u_2 \in U_2. \tag{2.127b}$$

In such case, the pair $(u_1^*,\ u_2^*)$ can be defined as

$$u_1^* = -P_1x \quad \text{and} \quad u_2^* = -P_2x, \text{ where} \tag{2.128}$$

$$-\dot{P}_1 = P_1(-1-P_2) + (-1-P_2)^TP_1 - P_1^2 + 1 + 2P_2^2 + P_1\varepsilon^2, \quad P_1(1) = 0,$$
$$-\dot{P}_2 = P_2(-1-P_1) + (-1-P_1)^TP_2 - P_2^2 + 1 + 2P_1^2 + P_1\varepsilon^2, \quad P_2(1) = 0.$$

which imply

$$\dot{P}_1 = P_1^2 + 2P_1P_2 + 2P_1 - 2P_2^2 - 1 + P_1\varepsilon^2, \quad P_1(1) = 0, \tag{2.129a}$$

$$\dot{P}_2 = P_2^2 + 2P_1P_2 + 2P_2 - 2P_1^2 - 1 + P_1\varepsilon^2, \quad P_2(1) = 0. \tag{2.129b}$$

Solving the system (2.129) for P_1 and P_2 we can find the optimal control from (2.128). The state trajectory can be depicted by Fig. 2.3 and can be detected by the following equation:

$$x(t+h) = x(t) + h(-1 - P_1 - P_2)x(t) + \varepsilon x(t)\sqrt{h}N(0,1), \quad x(0) = 1, \tag{2.130}$$

where $0 < h < 1$ is a small step size of time.

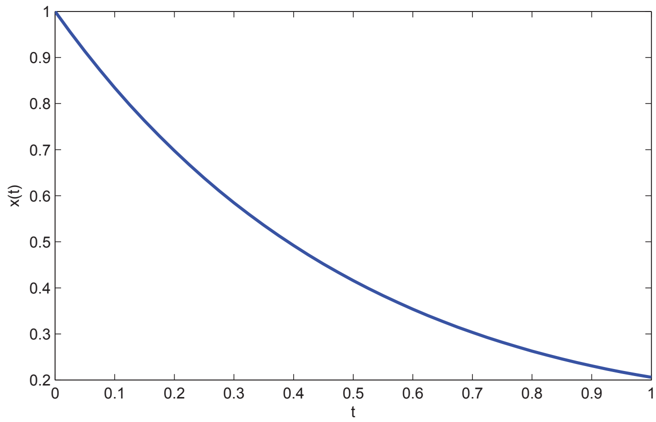


Fig. 2.3: Finite-horizon state trajectory.

Infinite-horizon case:

Let us consider the following system matrices of the system (2.115)–(2.116) for a two-player Nash equilibrium problem:

$$A = \begin{bmatrix} -0.52 & 1.12 & 0 \\ 0 & -0.24 & 1 \\ 0.23 & 0.85 & -0.16 \end{bmatrix}, \; A_p = 0.1A, \; x(0) = \begin{bmatrix} 1 \\ 0.5 \\ -0.6 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0.15 \\ 0.12 \\ 3.55 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0.23 \\ -0.52 \\ 0.28 \end{bmatrix},$$

$$Q_1 = \mathbf{diag}\begin{pmatrix} 1 & 1.5 & 2.1 \end{pmatrix}, \; Q_2 = \mathbf{diag}\begin{pmatrix} 1.2 & 1.1 & 3.1 \end{pmatrix},$$

$$R_{11} = 1.9, \quad R_{12} = 2.5, \quad R_{21} = 2.7, \quad R_{22} = 3.5.$$

Applying Lemma 2.3, we can obtain the following results:

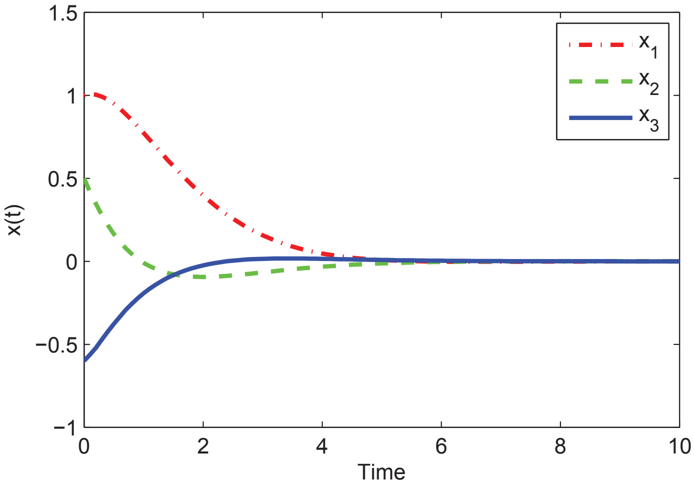


Fig. 2.4: Trajectory of the state.

$$P_1 = \begin{bmatrix} 8.4141e-1 & 6.2288e-1 & 1.0255e-1 \\ 6.2288e-1 & 2.2012 & 5.0173e-1 \\ 1.0255e-1 & 5.0173e-1 & 6.3970e-1 \end{bmatrix},$$
$$P_2 = \begin{bmatrix} 1.0076 & 8.1692e-1 & 1.4292e-1 \\ 8.1692e-1 & 2.5452 & 6.2007e-1 \\ 1.4292e-1 & 6.2007e-1 & 9.0459e-1 \end{bmatrix},$$
$$K_1 = \begin{bmatrix} -2.9737e-1 & -1.1256 & -1.2350 \end{bmatrix},$$
$$K_2 = \begin{bmatrix} 4.3723e-2 & 2.7486e-1 & 1.0366e-2 \end{bmatrix}.$$

It can be observed that the Lyapunov iterative algorithm converges to the exact solution with an accuracy of 10^{-13} after 14 iterations. Fig. 2.4 shows the response of the system with a state trajectory. It shows that the state variables $x(k)$ can stabilize the given system, which implies that the proposed method is very useful and reliable.

2.4 Disturbance attenuation problems

Consider the following stochastic linear system [Zhang and Chen (2004)]:

$$\begin{cases} dx(t) = [Ax(t) + B_2 u(t) + B_1 v(t)]\, dt + A_p x(t) dw(t), \\ z(t) = \begin{bmatrix} Cx(t) \\ Du(t) \end{bmatrix}, \end{cases} \tag{2.131}$$

where $x(0) = 0$ and $D^T D = I$. In (2.131), $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^{n_u}$ is the control input, $v(t) \in \mathbb{R}^{n_v}$ is the disturbance, $w(t) \in \mathbb{R}$ is a one-dimensional wiener process and $z(t) \in \mathbb{R}^{n_z}$ is the controlled output.

We want to minimize H_2 performance of the output by controlling $u(t)$ so that the effect of disturbance $v(t)$ will be eliminated under H_∞ - constraint. Moreover, minimize the desired cost function when worst-case disturbance $v^*(t)$ is imposed.

For any disturbance attenuation $\gamma > 0$, we need to find a state feedback control $u^*(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ such that,

(i) For $v \neq 0$, the perturbation operator

$$\begin{aligned}\|\mathscr{L}\|_\infty &= \sup_{\substack{v(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x_0 = 0}} \frac{\|z\|}{\|v\|}, \\ &= \sup_{\substack{v(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x_0 = 0}} \frac{\left\{\mathbb{E}\left[\int_0^\infty (x^T C^T C x + u^T u) dt\right]\right\}^{1/2}}{\left\{\mathbb{E}\left[\int_0^\infty v^T v\ dt\right]\right\}^{1/2}} < \gamma. \end{aligned} \tag{2.132}$$

(ii) For $v = 0$, $u^*(t)$ stabilizes system (2.131) internally, i.e, $\lim_{t\to\infty} \mathbb{E}[x^T(t)x(t)] = 0$.

(iii) For $v(t) = v^*(t)$, worst-case disturbance, if exist, where

$$v^*(t) = \arg\min_v J_1(x_0,\ u^*,\ v), \quad \forall x_0 \in \mathbb{R}^n, \tag{2.133}$$

with

$$J_1(x_0,\ u^*,\ v) = \mathbb{E}\left[\int_0^\infty (\gamma^2 \|v\|^2 - \|z\|^2) dt\right], \tag{2.134}$$

is applied to the system (2.131) $u^*(t)$ minimizes the cost functional,

$$J_2(x_0,\ u,\ v^*) = \|z\|^2 = \mathbb{E}\left[\int_0^\infty (x^T C^T C x + u^T u) dt\right]. \tag{2.135}$$

Equivalently condition (i), next theorem will also show that if $J_1(x_0,\ u^*,\ v^*) \geq 0$, then u^* is a solution to the stochastic H_2/H_∞ control. If an admissible control $u(t)$ satisfies the condition (i) and (ii), then $u(t)$ is called a solution under H_∞-constraint.

The infinite horizon stochastic H_2/H_∞ control is associated with the two-player, nonzero-sum Nash equilibrium strategies $(u^*,\ v^*)$ defined by,

$$J_1(x_0,\ u^*,\ v^*) \leq J_1(x_0,\ u^*,\ v), \tag{2.136}$$

$$J_2(x_0,\ u^*,\ v^*) \leq J_2(x_0,\ u,\ v^*). \tag{2.137}$$

If the previous $(u^*,\ v^*)$ exists, then we say that the infinite horizon H_2/H_∞ control admits a pair of solutions.

Theorem 2.8. *For* (2.131)*, suppose that the coupled AREs,*

$$P_1(A-B_2B_2^TP_2)+(A-B_2B_2^TP_2)^TP_1+A_p^TP_1A_p=\tilde{A}_2^T\tilde{A}_2, \tag{2.138}$$
$$P_2(A-B_2B_2^TP_2-\gamma^{-2}B_1B_1^TP_1)+(A-B_2B_2^TP_2-\gamma^{-2}B_1B_1^TP_1)^TP_2+A_p^TP_2A_p=-\tilde{A}_3^T\tilde{A}_3, \tag{2.139}$$

have a pair of solution $(P_1\le 0, P_2\ge 0)$*, where* $\tilde{A}_2=\begin{bmatrix} C \\ \gamma^{-1}B_1^TP_1 \\ B_2^TP_2\end{bmatrix}$ *and* $\tilde{A}_3=\begin{bmatrix} C \\ B_2^TP_2\end{bmatrix}$*. If* $[A,A_p|C]$ *and* $[A-\gamma^{-2}B_1B_1^TP_1,A_p|C]$ *are exactly observable, then the stochastic* H_2/H_∞ *control problem admits a pair of solutions*

$$u^*(t)=-B_2^TP_2x(t), \tag{2.140}$$
$$v^*(t)=-\gamma^{-2}B_1^TP_1x(t). \tag{2.141}$$

Proof. To prove Theorem 2.8, we have to prove the following claims:

(i) $(u^*,\ v^*)\in\mathscr{L}_{\mathscr{F}}^2(\mathbb{R}_+,\ \mathbb{R}^{n_u})\times\mathscr{L}_{\mathscr{F}}^2(\mathbb{R}_+,\ \mathbb{R}^{n_v})$ and $(A-B_2B_2^TP_2,A_p)$ is stable,

(ii) $\|\mathscr{L}\|_\infty<\gamma$ and

(iii) u^* minimizes the output energy $\|z\|_2^2$ when v^* applied in (2.131), i.e.,

$$u^*=\arg\min_u J_2(x_0,\ u,\ v^*),\qquad \forall u\in\mathscr{L}_{\mathscr{F}}^2(\mathbb{R}_+,\ \mathbb{R}_u^n).$$

(i) By Lemma 3 of [Zhang and Chen (2004)], $[A-B_2B_2^TP_2-\gamma^{-2}B_1B_1^TP_1,\ A_p|\tilde{A}_3]$ is exactly observable. So form Lemma 1 of [Zhang and Chen (2004)], (2.131) yields $(A-B_2B_2^TP_2-\gamma^{-2}B_1B_1^TP_1,A_p)$ being stable. Hence, $(u^*,\ v^*)\in\mathscr{L}_{\mathscr{F}}^2(\mathbb{R}_+,\ \mathbb{R}^{n_u})\times\mathscr{L}_{\mathscr{F}}^2(\mathbb{R}_+,\ \mathbb{R}^{n_v})$.

Second, form Lemmas 1 and 3 of [Zhang and Chen (2004)], (2.138) yields $(A-B_2^TP_2P_2,A_p)$ is stable, i.e., (2.131) is internally stabilizable by $u(t)=u^*(t)=-B_2^TP_2x(t)$. So, we can write [Hinrichsen and Pritchard (1998)],

$$P_1(A-B_2B_2^TP_2)+(A-B_2B_2^TP_2)^TP_1+A_p^TP_1A_p-P_2B_2B_2^TP_2-\gamma^{-2}P_1B_1B_1^TP_1-C^TC=0, \tag{2.142}$$

(ii) Substituting $u(t)=u^*(t)=-B_2^TP_2x(t)$ into (2.131) gives

$$\begin{cases} dx(t)=\left\{(A-B_2B_2^TP_2)x(t)+B_1v(t)\right\}dt+A_px(t)dw(t), \\ z(t)=\begin{bmatrix} Cx(t) \\ -DB_2^TP_2x(t)\end{bmatrix}, \end{cases} \tag{2.143}$$

where $x(0) = x_0$. Applying Ito's formula to (2.143) and considering (2.138), we have

$$\mathbb{E}\left[\int_0^\infty d(x^T P_1 x)\right] = \mathbb{E}\left[\int_0^\infty \left\{\left((A - B_2 B_2^T P_2)x + B_1 v\right)(P_1 x + x^T P_1) + x^T A_p^T P_1 A_p x\right\} dt\right]$$

$$\begin{aligned}
\text{or, } -x_0^T P_1 x_0 = \mathbb{E}\Big[\int_0^\infty \Big\{ & x^T (P_1(A - B_2 B_2^T P_2) + (A - B_2 B_2^T P_2)^T P_2 + A_p^T P_1 A_p)x \\
& + v^T B_1^T P_1 x + x^T P_1 B_1 v\Big\} dt\Big] \\
= \mathbb{E} & \left[\int_0^\infty \{x^T \tilde{A}_2^T \tilde{A}_2 x + v^T B_1^T P_1 x + x^T P_1 B_1 v\} dt\right] \\
= \mathbb{E} & \left[\int_0^\infty \{x^T (C^T C + \gamma^{-2} P_1 B_1 B_1^T P_1 + P_2 B_2 B_2^T P_2)x + v^T B_1^T P_1 x + x^T P_1 B_1 v\} dt\right] \\
= \mathbb{E} & \left[\int_0^\infty \left\{z^T z + \gamma^2 {v^*}^T v^* + v^T B_1^T P_1 x + x^T P_1 B_1 v\right\} dt\right] \\
& [\text{suppose, } v^*(t) = -\gamma^{-2} B_1^T P_1 x(t)]
\end{aligned}$$

$$\begin{aligned}
\text{or, } \mathbb{E}\left[\int_0^\infty \{\gamma^2 v^T v - z^T z\} dt\right] &= x_0^T P_1 x_0 + \mathbb{E}\left[\int_0^\infty \left\{\gamma^2 v^T v + \gamma^2 {v^*}^T v^* - \gamma^2 v^T v^* - \gamma^2 {v^*}^T v\right\} dt\right] \\
&= x_0^T P_1 x_0 + \gamma^2 \mathbb{E}\left[\int_0^\infty (v - v^*)^T (v - v^*) dt\right] \qquad (2.144)
\end{aligned}$$

So

$$\begin{aligned}
J_1(x_0,\, u^*,\, v) &= \mathbb{E}\left[\int_0^\infty \{\gamma^2 v^T v - z^T z\} dt\right] \\
&= x_0^T P_1 x_0 + \gamma^2 \mathbb{E}\left[\int_0^\infty (v - v^*)^T (v - v^*) dt\right] \geq J_1(x_0,\, u^*,\, v^*) = x_0^T P_1 x_0.
\end{aligned} \qquad (2.145)$$

Now, if we define an operator $\mathscr{L}_1 v = v - v^*$, then form (2.145) we have (for $x(0) = x_0 = 0$):

$$J_1(x_0,\, u^*,\, v) = \gamma \|v\|^2 - \|z\|^2 = \gamma^2 \|\mathscr{L}_1 v\|^2 \geq \varepsilon \|v\|^2 > 0$$

, for some $\varepsilon > 0$, which yields $\|\mathscr{L}\|_\infty < \gamma$.

(iii) Finally, when worst-case disturbance $v = v^*(t) = -\gamma^{-2} B_1^T P_1 x(t)$ is applied to (2.131), we have

$$\begin{cases} dx(t) = \left\{(A - \gamma^{-2} B_1 B_1^T P_1)x(t) + B_2 u(t)\right\} dt + A_p x(t) dw(t) \\ z(t) = \begin{bmatrix} Cx(t) \\ Du(t) \end{bmatrix}, \end{cases} \qquad (2.146)$$

where $x(0) = x_0$. Now the H_2 optimization problem becomes a standard stochastic LQ optimal control problem, so we can write [Rami and Zhou (2000)]

$$P_2(A - \gamma^{-2} B_1 B_1^T P_1) + (A - \gamma^{-2} B_1 B_1^T P_1)^T P_2 + A_p^T P_1 A_p - P_2 B_2 B_2^T P_2 + C^T C = 0, \quad (2.147)$$

which is the same as (2.139). Applying Ito's formula in (2.146) considering (2.147) we get,

$$\mathbb{E}\left[\int_0^\infty d(x^T P_2 x)\right] = \mathbb{E}\left[\int_0^\infty \left\{\left((A-\gamma^{-2}B_1B_1^TP_1)x + B_2u\right)(P_2x + x^TP_2) + x^TA_p^TP_2A_px\right\} dt\right]$$

$$\text{or, } -x_0^TP_2x_0 = \mathbb{E}\left[\int_0^\infty \left\{x^TP_2B_2B_2^TP_2x - x^TC^TCx + u^TB_2^TP_2x + x^TP_2B_2u\right\} dt\right] \text{ [by (2.147)]}$$

$$\text{or, } \mathbb{E}\left[\int_0^\infty \left\{x^TC^TCx + u^Tu\right\} dt\right]$$

$$= x_0^TP_2x_0 + \mathbb{E}\left[\int_0^\infty \left\{u^Tu + {u^*}^Tu^* + u^Tu^* - u^Tu^* - {u^*}^Tu\right\} dt\right]$$

$$\text{or, } J_2(x_0,\ u,\ v^*) = x_0^TP_2x_0 + \mathbb{E}\left[\int_0^\infty (u-u^*)^T(u-u^*)dt\right]. \tag{2.148}$$

If we put $u = u^*$, then form (2.148) we get

$$J_2(x_0,\ u,\ v^*) \geq J_2(x_0,\ u^*,\ v^*) = x_0^TP_2x_0. \tag{2.149}$$

It can be shown that $[A-\gamma^{-2}B_1^TP_1, A_p|C]$ is exactly observable and $[A-B_2B_2^TP_2-\gamma^{-2}B_1^TP_1, A_p|C]$ is stochastically detectable. So, the maximal solution of AREs (2.138) and (2.139) can be written as $(P_1 \leq 0,\ P_2 \geq 0)$. □

Chapter 3

H_∞-Constrained Incentive Stackelberg Game for Discrete-Time Systems with Multiple Non-cooperative Followers

This chapter is based on a previously published article [Ahmed and Mukaidani (2016)].

3.1 Introduction

Stackelberg leadership model is a hierarchical strategy involving the first movement of the leader and then the consequent movement of followers. With several control problems, dynamic games for both continuous- and discrete-time systems have been extensively studied (see e.g. [Başar and Olsder (1999)] and references therein). Recently, due to the growth of interest in multi-agent and cooperative systems, the theoretical game problem and the applications have been widely investigated. The interest in multi-agent cooperative systems with theoretical game problems and applications is increasing. For example, a new class of multi-agent discrete-time dynamic games are demonstrated in terms of the solutions of the discrete-time coupled Hamilton Jacobi equations [Abouheaf et al. (2013)]. In [Shen (2004)], a non-cooperative game with Nash equilibrium state feedback control has been considered. Subsequently, Riccati design techniques and neural adaptive design techniques for cooperative control of multi-agent systems with unknown dynamics has been established in [Lewis et al. (2013)].

The open and closed-loop Stackelberg games are commonly used in dynamic non-cooperative games and the hierarchical decision making problems [Başar and Olsder (1999)], [Medanic (1978)]. The basic feature of the Stackelberg game is the leader determines his strategy ahead and the followers optimize their own cost subject to the leader's announcement. At last, the leader optimize his cost considering the optimized followers' constraints. Also, incentive Stackelberg strategy an extensive idea in which the leader can achieve his team-optimal solution in a Stackelberg game. Over the past 40 years, the incentive Stackelberg strategy was studied intensively (see e.g. [Ho et al. (1982), Basar and Selbuz (1979), Basar and Olsder (1980), Zheng and Basar (1982), Zheng et al. (1984)] and references therein). In [Li et al. (2002)], the team-optimal state feedback incentive Stackelberg strategy of discrete-time two-player nonzero-sum dynamic games characterized by linear state dynamics and quadratic cost functionals was developed. However, the deterministic disturbance is not taken into account in these literatures. To the best of our knowledge, such perspective is lacking in the literature in view of the case of the existence of the external disturbance in hierarchical control strategy. Therefore, the incentive Stackelberg game for such systems seem to be even more challenging.

In this Chapter, the incentive Stackelberg game for a discrete-time system with multiple followers under H_∞ constraint is considered. We discuss only two-level hierarchical games with one leader and many non-cooperative followers. In our work, the conditions for the existence of the leader's team-optimal solution under the H_∞ constraint are derived based on the existing results in [Zhang et al. (2007)]. It is shown that a solution can be found by solving a set of cross-coupled backward difference Riccati equations (CCBDREs). Moreover, the followers' strategies are established in such a way that satisfies the leader's team-optimal solution. Furthermore, we discuss the infinite-horizon case and propose a numerical algorithm to obtain a solution set of the coupled algebraic Riccati equations. A numerical example demonstrates the efficiency of the proposed methodology.

Notation: The notations used in this Chapter are fairly standard. I_n denotes the $n \times n$ identity matrix. **block diag** denotes the block diagonal matrix. $[\cdot]$ denotes the expectation operator. $\boldsymbol{Y} = \{y(k) : y(k) \in \mathbb{R}^n\}_{0 \le k \le T} = \{y(0), y(1), \dots, y(T_f)\}$ denotes the finite sequences. The l^2-norm of $y(k) \in l^2(\mathbf{N}_{T_f}, \mathbb{R}^n)$ is defined by $\|y(k)\|^2_{l^2(\mathbf{N}_{T_f}, \mathbb{R}^n)} := \sum_{k=0}^{T_f} [\|y(k)\|^2]$, where $\mathbf{N}_{T_f} := \{0, 1, \dots, T_f\}$.

3.2 Preliminary results

Consider the linear discrete-time system,

$$x(k+1) = A(k)x(k) + B(k)u(k), \tag{3.1a}$$

$$x(0) = x^0, \quad k = 0, 1, \dots, T_f, \tag{3.1b}$$

where $x(k) \in \mathbb{R}^n$ represents the state vector, $u(k) \in \mathbb{R}^m$ denotes the control input, $A(k)$ and $B(k)$ are assumed to be matrix-valued functions of suitable dimensions.

Let us define the cost functional,

$$J(x^0,\ u) := \frac{1}{2}\sum_{k=0}^{T_f}[x^T(k)Q(k)x(k) + 2x^T(k)S(k)u(k) + u^T(k)R(k)u(k)], \tag{3.2}$$

where $Q(k) = Q^T(k) \geq 0$, $R(k) = R^T(k) > 0$, $Q(k) - S^T(k)[R(k)]^{-1}S(k) > 0$ and $0 < T < \infty$. By using the similar technique of [Zhang et al. (2008)] and [Rami et al. (2002)] to find an admissible control of the above system, we can derive the following lemma3:

Lemma 3.1. *Suppose that the following backward difference Riccati equation (BDRE) has solution(s):*

$$\begin{aligned} X(k) =& A_S^T(k)X(k+1)A_S(k) - A_S^T(k)X(k+1)V(k)X(k+1)A_S(k) + Q(k) \\ & - S^T(k)[R(k)]^{-1}S(k), \quad X(T+1) = 0, \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} &A_S(k) := A(k) - B(k)[R(k)]^{-1}S(k), \\ &V(k) := B(k)[\hat{R}(k)]^{-1}B^T(k), \\ &\hat{R} := R(k) + B^T(k)X(k+1)B(k), \\ &Q(k) - S^T(k)[R(k)]^{-1}S(k) \geq 0. \end{aligned}$$

Then, the optimal state feedback control is given by

$$u^*(k) = K^*(k)x(k) \tag{3.4}$$

$$= -[\hat{R}(k)]^{-1}\big(B^T(k)X(k+1)A_S(k) + [R(k)]^{-1}S(k)\big)x(k). \tag{3.5}$$

In contrast to [Zhang et al. (2008)] and [Rami et al. (2002)], there exist a cross-coupling term $2x^T(k)S(k)u(k)$ *in the cost functional* (3.1) *in a deterministic system.*

Proof. Let us consider the Hamilton as follows:

$$\begin{aligned} H(k) =& x^T(k)Q(k)x(k) + 2x^T(k)S(k)u(k) + u^T(k)R(k)u(k) + V(k+1) \\ =& x^T(k)Q(k)x(k) + 2x^T(k)S(k)u(k) + u^T(k)R(k)u(k) \\ & + x^T(k+1)X(k+1)x(k+1), \end{aligned} \tag{3.6}$$

where $V(k)$ is a quadratic function as follows:

$$V(k) = \frac{1}{2}x^T(k)X(k)x(k), \quad V(T_f) = 0, \tag{3.7}$$

with the symmetric positive semi-definite matrix $X(k)$.

Using equation (3.1a), the following result can be found form (3.6):

$$
\begin{aligned}
H(k) =& x^T(k)Q(k)x(k) + 2x^T(k)S(k)u(k) + u^T(k)R(k)u(k) \\
& + [A(k)x(k) + B(k)u(k)]^T X(k+1)[A(k)x(k) + B(k)u(k)], \\
=& x^T(k)[Q(k) + A^T(k)X(k+1)A(k)]x(k) \\
& + 2x^T(k)[S(k) + A^T(k)X(k+1)B(k)]u(k) + u^T(k)[R(k) + B^T(k)X(k+1)B(k)]u(k),
\end{aligned} \tag{3.8}
$$

To minimize the right hand side of (3.8) with respect to the control input $u(k)$ we obtain the following state feedback optimal control scheme:

$$
\begin{aligned}
u^*(k) &= -[R(k) + B^T(k)X(k+1)B(k)]^{-1}[S(k) + A^T(k)X(k+1)B(k)]^T x(k), \\
&= -[\hat{R}(k)]^{-1}\big(B^T(k)X(k+1)A_S(k) + [R(k)]^{-1}S(k)\big)x(k),
\end{aligned} \tag{3.9}
$$

where

$$
\begin{aligned}
A_S(k) &:= A(k) - B(k)[R(k)]^{-1}S(k), \\
\hat{R} &:= R(k) + B^T(k)X(k+1)B(k).
\end{aligned}
$$

Comparing right hand sides of (3.49) and (3.8), the following backward difference Riccati equation (BDRE) can be derived using (3.3):

$$
\begin{aligned}
X(k) =& A_S^T(k)X(k+1)A_S(k) - A_S^T(k)X(k+1)V(k)X(k+1)A_S(k) + Q(k) \\
& - S^T(k)[R(k)]^{-1}S(k), \quad X(T+1) = 0.
\end{aligned} \tag{3.10}
$$

□

On the other hand, consider the following discrete-time system.

$$
x(k+1) = A(k)x(k) + D(k)v(k), \tag{3.11a}
$$

$$
z(k) = C(k)x(k), \quad x(0) = x_0, \quad k = 0,\ 1,\ \ldots T_f, \tag{3.11b}
$$

where $v(k) \in \mathbb{R}^{n_v}$ represents the external disturbance. $z(k) \in \mathbb{R}^{n_z}$ represents the controlled output. The following definition is the counterpart of the deterministic case of the existing results in [Zhang et al. (2007), Zhang et al. (2008)].

Definition 3.1. *In system* (3.11)*, if the disturbance input $v(k) \in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})$ and the controlled output $z(k) \in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})$, then the perturbed operator*

$$
L_{T_f} := l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}) \to l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z}) \tag{3.12}
$$

is defined by

$$L_{T_f}v(k) := Cx(k,\ 0,\ v),\ \forall v(k) \in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}), \tag{3.13}$$

with its norm

$$\|L_{T_f}\| := \sup_{\substack{v(k)\in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}),\\ v(k)\neq 0,\ x^0=0}} \frac{\|z(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})}}{\|v(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})}}. \tag{3.14}$$

The following lemma can be viewed as the deterministic version of [Zhang et al. (2007), Zhang et al. (2008)].

Lemma 3.2. *For the discrete time system* (3.11), $\|L_{T_f}\| < \gamma$ *for given* $\gamma > 0$ *if and only if there exists a unique solution* $Y(k) \leq 0$ *to the following matrix difference equation.*

$$\begin{aligned} Y(k) = A^T(k)Y(k+1)A(k) - A^T(k)Y(k+1)U(k)Y(k+1)A(k) \\ - C^T(k)C(k), \qquad Y(T+1) = 0, \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} U(k) &:= D(k)[T_\gamma(k)]^{-1}D^T(k), \\ T_\gamma(k) &:= \gamma^2 I_{n_v} + D^T(k)Y(k+1)D(k). \end{aligned}$$

In this case, worst-case disturbance is given by

$$v^*(k) = F^*_\gamma(k)x(k) = -[T_\gamma(k)]^{-1}D^T(k)Y(k+1)A(k)x(k). \tag{3.16}$$

Proof. The corresponding cost function can be defined as:

$$J_v(x_0,\ v) := \sum_{k=0}^{T_f}[\gamma^2 v^T(k)v(k) - x^T(k)C^T(k)C(k)x(k)]. \tag{3.17}$$

Let us consider the Hamilton as follows:

$$\begin{aligned} H(k) =& \gamma^2 v^T(k)v(k) - x^T(k)C^T(k)C(k)x(k) + V(k+1) \\ =& \gamma^2 v^T(k)v(k) - x^T(k)C^T(k)C(k)x(k) + x^T(k+1)Y(k+1)x(k+1), \end{aligned} \tag{3.18}$$

where $V(k)$ is a quadratic function as follows:

$$V(k) = \frac{1}{2}x^T(k)Y(k)x(k), \quad V(T_f) = 0, \tag{3.19}$$

with the symmetric positive semi-definite matrix $Y(k)$.

Using equation (3.11a), the following result can be found form (3.18):

$$
\begin{aligned}
H(k) =& \gamma^2 v^T(k)v(k) - x^T(k)C^T(k)C(k)x(k) \\
&+ [A(k)x(k) + D(k)v(k)]^T Y(k+1)[A(k)x(k) + D(k)v(k)], \\
=& x^T(k)[A^T(k)Y(k+1)A(k) - C^T(k)C(k)]x(k) \\
&+ 2x^T(k)[A^T(k)Y(k+1)D(k)]v(k) + v^T(k)[\gamma^2 I_{n_v} + D^T(k)Y(k+1)D(k)]v(k),
\end{aligned}
\tag{3.20}
$$

To minimize the right hand side of (3.20) with respect to the disturbance input $v^*(k)$ we obtain the following state feedback worst-case disturbance:

$$
\begin{aligned}
v^*(k) &= -[\gamma^2 I_{n_v} + D^T(k)Y(k+1)D(k)]^{-1}[A^T(k)Y(k+1)D(k)]^T x(k), \\
&= -[T_\gamma(k)]^{-1} D^T(k)Y(k+1)A(k)x(k)
\end{aligned}
\tag{3.21}
$$

where

$$
T_\gamma(k) := \gamma^2 I_{n_v} + D^T(k)Y(k+1)D(k).
$$

Comparing right hand sides of (3.19) and (3.20), the following backward difference Riccati equation (BDRE) can be derived using (3.56):

$$
\begin{aligned}
Y(k) = &A^T(k)Y(k+1)A(k) - A^T(k)Y(k+1)U(k)Y(k+1)A(k) \\
&- C^T(k)C(k), \qquad Y(T+1) = 0,
\end{aligned}
\tag{3.22}
$$

where

$$
U(k) := D(k)[T_\gamma(k)]^{-1} D^T(k).
$$

□

3.3 H_∞-constrained incentive Stackelberg game

3.3.1 Problem formulation

Consider a linear discrete-time system involving multiple followers defined by

$$
x(k+1) = A(k)x(k) + \sum_{j=1}^{N} B_{0j}(k)u_{0j}(k) + \sum_{j=1}^{N} B_j(k)u_j(k) + D(k)v(k), \tag{3.23a}
$$

$$
z(k) = \begin{bmatrix} C(k)x(k) \\ G_0(k)\boldsymbol{u}_0(k) \\ G_1(k)u_1(k) \\ \vdots \\ G_N(k)u_N(k) \end{bmatrix}, \tag{3.23b}
$$

where $x(0)=x_0,\ G_i^T(k)G_i(k)=I_{m_i},\ u_i(k)\in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_i}),\ i=1,\ \dots\ ,N$ represents the i-th follower's control input. It should be noted that $i=0$ represents the leader's control input,

$$\boldsymbol{u}_0(k)=\left[\begin{array}{ccc} u_{01}^T(k) & \cdots & u_{0N}^T(k) \end{array}\right]^T,$$

and each i-th control input $u_{0i},\ i=1,\ \dots\ ,N$ is applied for each i-th follower. Define the linear quadratic cost functionals as follows:

$$\begin{aligned} & J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots\ ,u_N,\ v) \\ & := \frac{1}{2}\sum_{k=0}^{T_f}\left[x^T(k)Q_0(k)x(k)+\sum_{j=1}^{N}\left\{u_{0j}^T(k)R_{00j}(k)u_{0j}(k)+u_j^T(k)R_{0j}(k)u_j(k)\right\}\right], \end{aligned} \tag{3.24a}$$

$$\begin{aligned} & J_i(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots\ ,u_N,\ v) \\ & := \sum_{k=0}^{T_f}\left[x^T(k)Q_i(k)x(k)+u_{0i}^T(k)R_{0ii}(k)u_{0i}(k)+u_i^T(k)R_{ii}(k)u_i(k)\right], \end{aligned} \tag{3.24b}$$

where $Q_i(k)=Q_i^T(k)\geq 0,\ R_{00i}(k)=R_{00i}^T(k)>0,\ R_{0i}(k)=R_{0i}^T(k)\geq 0,\ R_{0ii}(k)=R_{0ii}^T(k)\geq 0,\ R_{ii}(k)=R_{ii}^T(k)>0,\ i=1,\ \dots\ ,N$. Furthermore, for given a disturbance attenuation level $\gamma>0$, define the performance function

$$J_\gamma(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots\ ,u_N,\ v) := \sum_{k=0}^{T_f}[\gamma^2\|v(k)\|^2-\|z(k)\|^2]. \tag{3.25}$$

It should be noted that throughout the Chapter, each player only has access to perfect state information and the following state feedback form defined by the space of admissible strategies $\Gamma_i,\ i=0,\ 1,\ \dots\ ,N$ is considered.

$$\left.\begin{array}{l} u_{0i}(k)=u_{0i}\big(k,\ x(k),\ x(0)\big) \\ u_i(k)=u_i\big(k,\ x(k),\ x(0)\big) \end{array}\right\},\quad i=1,\ \dots\ ,N. \tag{3.26}$$

The finite horizon H_∞-constrained incentive Stackelberg game with multiple non-cooperative followers is given below:

Given the disturbance attenuation level $\gamma>0,\ 0<T<\infty$, find (if possible) strategies $u_{0i}^*(k)\in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_{0i}}),\ u_i^*(k)\in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_i}),\ i=1,\ \dots\ ,N$ such that

i) *for the worst-case disturbance $v^*(k)\in l^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})$, the following inequalities hold:*

$$J_0(x_0,\ \boldsymbol{u}_0^*,\ u_1^*,\ \dots\ ,u_N^*,\ v^*)=\min_{(\boldsymbol{u}_0,\ u_1,\ \dots\ ,u_N)} J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots\ ,u_N,\ v^*), \tag{3.27a}$$

$$J_i(x_0,\ u_{01}^*,\ \dots\ ,u_{0N}^*,\ u_1^*,\ \dots\ ,u_N^*,\ v^*)\leq J_i(x_0,\ u_{0(-i)}^*,\ u_{-i}^*,\ v^*), \tag{3.27b}$$

where

$$
\begin{aligned}
&u_{0(-i)}^* := (u_{01}^*,\ \dots\ ,u_{0(i-1)}^*,\ u_{0i},\ u_{0(i+1)}^*,\ \dots\ ,u_{0N}^*),\\
&u_{-i}^* := (u_1^*,\ \dots\ ,u_{i-1}^*,\ u_i,\ u_{i+1}^*,\ \dots\ ,u_N^*),\\
&u_{0j} := u_{0j}(u_j),\ u_{0j}^* = u_{0j}(u_j^*), \quad j = 1,\ \dots\ ,N,\\
&u_j := u_j(u_j),\ u_j^* = u_j(u_j^*), \quad j = 1,\ \dots\ ,N.
\end{aligned}
$$

ii) *The norm of the perturbed operator mentioned in* (3.14) *and the disturbance attenuation level are related as*

$$
\|L_{T_f}\| < \gamma, \tag{3.28}
$$

where $\|z(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^n)}$ *and* $\|v(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^n)}$ *in* (3.14) *are defined as*

$$
\begin{aligned}
\|z(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^n)} &:= \sum_{k=0}^{T_f}\left[\|C(k)x(k)\|^2 + \|\boldsymbol{u}_0^*(k)\|^2 + \sum_{j=1}^{N}\|u_j^*(k)\|^2\right]\\
\|v(k)\|^2_{l^2(\mathbf{N}_{T_f},\ \mathbb{R}^n)} &:= \sum_{k=0}^{T_f}[\|v(k)\|^2].
\end{aligned}
$$

It should be noted that a strategy pair $(\boldsymbol{u}_0^*,\ u_1^*,\ \dots\ ,u_N^*)$ is called a team-optimal strategy pair for the leader [Başar and Olsder (1999)]. The problem is that find sufficient conditions such that Stackelberg strategy achieves a team-optimal value for J_0. Furthermore, the condition of (3.27b) is called Nash equilibrium condition.

3.3.2 Main results

First, the team optimization problem is solved by using the standard linear quadratic (LQ) control under the worst disturbance. Let us consider the following LQ control problem.

$$
x(k+1) = A_v(k)x(k) + B_c(k)u_c(k), \tag{3.29a}
$$

$$
J_0(x_0,\ u_c) := \frac{1}{2}\sum_{k=0}^{T_f}[x^T(k)Q_0(k)x(k) + u_c^T(k)R_c(k)u_c(k)], \tag{3.29b}
$$

where

$$
\begin{aligned}
&A_v(k) := A(k) + D(k)F_\gamma(k),\\
&v^*(k) := F_\gamma(k)x(k),\\
&u_c(k) := \begin{bmatrix} \boldsymbol{u}_0^T(k) & u_1^T(k) & \cdots & u_N^T(k) \end{bmatrix}^T,
\end{aligned}
$$

$$
\begin{aligned}
&B_c(k) := \left[\ \boldsymbol{B}_0(k) \quad B_1(k) \quad \cdots \quad B_N(k)\ \right], \\
&\boldsymbol{B}_0(k) := \left[\ B_{01}(k) \quad \cdots \quad B_{0N}(k)\ \right], \\
&R_c(k) := \textbf{block diag}\left(\ \boldsymbol{R}_0(k) \quad R_{01}(k) \quad \cdots R_{0N}(k)\ \right), \\
&\boldsymbol{R}_0(k) := \textbf{block diag}\left(\ R_{001}(k) \quad \cdots R_{00N}(k)\ \right).
\end{aligned}
$$

According to Lemma 3.1, the team-optimal control can be written as

$$
\begin{aligned}
\bar{u}_c(k) &= \left[\ \bar{\boldsymbol{u}}_0^T(k) \quad \bar{u}_1^T(k) \quad \cdots \quad \bar{u}_N^T(k)\ \right]^T = K_c(k)x(k) = \begin{bmatrix} \boldsymbol{K}_0(k) \\ \boldsymbol{K}_1(k) \end{bmatrix} x(k) \\
&= -[\hat{R}_c(k)]^{-1} B_c^T(k) P(k+1) A_v(k) x(k), \qquad (3.30)
\end{aligned}
$$

where $P(k)$ satisfies the following BDRE

$$
\begin{aligned}
P(k) = A_v^T(k)P(k+1)A_v(k) - A_v^T(k)P(k+1)S_c(k)P(k+1)A_v(k) \\
+ Q_0(k), \qquad P(T+1) = 0, \qquad (3.31)
\end{aligned}
$$

with

$$
\begin{aligned}
&\boldsymbol{K}_0(k) := \left[\ K_{01}^T(k) \quad \cdots \quad K_{0N}^T(k)\ \right]^T, \\
&\boldsymbol{K}_1(k) := \left[\ K_1^T(k) \quad \cdots \quad K_N^T(k)\ \right]^T, \\
&S_c(k) := B_c(k)[\hat{R}_c(k)]^{-1} B_c^T(k), \\
&\hat{R}_c(k) := R_c(k) + B_c^T(k) P(k+1) B_c(k).
\end{aligned}
$$

Furthermore, the related team-optimal state response is given below.

$$
\bar{x}(k+1) = A_v(k)\bar{x}(k) + B_c(k)u_c(k) = [A_v(k) + B_c(k)K_c(k)]\bar{x}(k), \quad \bar{x}(0) = x_0. \qquad (3.32)
$$

On the other hand, by using Lemma 3.2, the H_∞ constraint condition can be obtained. Namely, suppose that the following BDRE has the solution set.

$$
\begin{aligned}
W(k) =& A_K^T(k)W(k+1)A_K(k) - A_K^T(k)W(k+1)U_W(k)W(k+1)A_K(k) \\
&- L_K(k), \quad W(T+1) = 0, \qquad (3.33)
\end{aligned}
$$

where

$$
\begin{aligned}
&A_K(k) := A(k) + B_c(k)K_c(k), \\
&U_W(k) := D(k)[T_{W\gamma}(k)]^{-1} D^T(k), \\
&T_{W\gamma}(k) := \gamma^2 I_{n_v} + D^T(k)W(k+1)D(k), \\
&L_K(k) := C^T(k)C(k) + \boldsymbol{K}_0^T(k)\boldsymbol{K}_0(k) + \boldsymbol{K}_1^T(k)\boldsymbol{K}_1(k).
\end{aligned}
$$

In this case, the worst-case disturbance is given by

$$v^*(k) = F_\gamma^*(k)x(k) = -[T_{W\gamma}(k)]^{-1}D^T(k)W(k+1)A_K(k)x(k). \tag{3.34}$$

It is assumed that the leader chooses the following incentive Stackelberg strategy:

$$u_{0i}^*(k) = \eta_{0i}(k)x(k) + \eta_{ii}(k)u_i(k),\ i = 1,\ 2,\ \dots\ ,N, \tag{3.35}$$

where $\eta_{0i}(k) \in \mathbb{R}^{m_{0i}\times n}$ and $\eta_{ii}(k) \in \mathbb{R}^{m_{0i}\times m_i}$ are strategy parameter matrices having the following relation:

$$\eta_{0i}(k) = K_{0i}(k) - \eta_{ii}(k)K_i(k),\ i = 1,\ 2,\ \dots\ ,N. \tag{3.36}$$

It should be ensured that $u_{0i}^*(k)$, $u_i^*(k)$ and $x^*(k)$ are satisfied the H_∞ constraint team-optimal Nash equilibrium. Hence for $i = 1,\ 2,\ \dots\ ,N$,

$$u_{0i}^*(k) = -[R_{00i}(k)]^{-1}B_{0i}^T P(k+1)x^*(k+1), \tag{3.37a}$$

$$u_i^*(k) = -[R_{0i}(k)]^{-1}B_i^T P(k+1)x^*(k+1). \tag{3.37b}$$

Second, the followers' optimization problem is solved. Consider the following cost functional,

$$\begin{aligned} &J_i(x_0,\ \boldsymbol{u}_0^*,\ u_1,\ \dots\ ,u_N,\ v) \\ &\quad := \frac{1}{2}\sum_{k=0}^{T_f}\left[x^T(k)Q_i(k)x(k) + u_{0i}^{*T}(k)R_{0ii}(k)u_{0i}^*(k) + u_i^T(k)R_{ii}(k)u_i(k)\right],\ i = 1,\ \dots\ ,N, \end{aligned} \tag{3.38}$$

where $\boldsymbol{u}_0^*(k) = \boldsymbol{u}_0^*(k,\ x(k),\ x(0))$ can be obtained by (3.35).
In order to establish the sufficient condition for optimality, the following Hamiltonian is defined.

$$\begin{aligned} H_i(u_i,\ \alpha_i) :=& \frac{1}{2}\left[x^T(k)Q_i(k)x(k) + u_{0i}^{*T}(k)R_{0ii}(k)u_{0i}^*(k) + u_i^T(k)R_{ii}(k)u_i(k)\right] \\ &+ \alpha_i^T(k+1)\Bigg[A_v(k)x(k) + B_{0i}(k)u_{0i}^*(k) + \sum_{j=1,\ j\neq i}^{N} B_{0j}(k)u_{0j}^*(k) \\ &+ B_i(k)u_i(k) + \sum_{j=1,\ j\neq i}^{N} B_j(k)u_j(k)\Bigg]. \end{aligned} \tag{3.39}$$

Hence we have,

$$\alpha_i(k) = \frac{\partial H_i(u_i,\ \alpha_i)}{\partial x(k)} = \tilde{Q}_i(k)x(k) + \tilde{A}_i^T(k)\alpha_i(k+1), \quad \alpha_i(T+1) = 0, \tag{3.40}$$

where

$$\tilde{A}_i(k) := A_v(k) + B_{0i}(k)\eta_{0i}(k) + \sum_{j=1,\ j\neq i}^{N} [B_{0j}(k)K_{0j}(k) + B_j(k)K_j(k)],$$
$$\tilde{Q}_i(k) := Q_i(k) + \eta_{0i}^T(k)R_{0ii}(k)\eta_{0i}(k).$$

Now, consider $\alpha_i(k) = P_i(k)x(k)$ then the following BDRE can be derived:

$$P_i(k) = \tilde{A}_i^T(k)P_i(k+1)\tilde{A}_i(k) - \tilde{X}_i^T(k)\tilde{Y}_i^{-1}(k)\tilde{X}_i(k) \tag{3.41}$$
$$+ \tilde{Q}_i(k), \quad P_i(T+1) = 0, \quad i = 1,\ 2,\ \ldots,N, \tag{3.42}$$

where

$$\tilde{X}_i(k) := \tilde{B}_i^T(k)P_i(k+1)\tilde{A}_i(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{0i}(k),$$
$$\tilde{Y}_i(k) := \tilde{R}_i(k) + \tilde{B}_i^T(k)P_i(k+1)\tilde{B}_i(k),$$
$$\tilde{B}_i(k) := B_i(k) + B_{0i}(k)\eta_{ii}(k),$$
$$\tilde{R}_i(k) := R_{ii}(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{ii}(k).$$

The followers' optimal strategy will be determined by

$$\frac{\partial H_i(u_i,\ \alpha_i)}{\partial u_i(k)} = R_{ii}(k)u_i^*(k) + B_i(k)^T\alpha_i(k+1) = 0, \tag{3.43}$$

which implies

$$u_i^*(k) = \tilde{K}_i(k)x(k) = [\tilde{Y}_i(k)]^{-1}\tilde{X}_i(k)x(k),\ i = 0,\ 1,\ \ldots,N. \tag{3.44}$$

Remark 3.1. *It should be noted that the incentive parameter $\eta_{ii}(k)$ can be uniquely determined if and only if $\tilde{Y}_i(k)$ is non-singular.*

3.4 Infinite horizon case

The infinite-horizon H_∞-constrained incentive Stackelberg game is studied in this section.

Definition 3.2. *[Zhang et al. (2008)] The following discrete-time system:*

$$\begin{cases} x(k+1) = Ax(k), \\ z(k) = Cx(k), \quad x(0) = x_0 \in \mathbb{R}^n, \quad k \in \mathbf{N}_{T_f}, \end{cases} \tag{3.45}$$

or $(A,\ C)$ is said to be exactly observable if $z(k) \equiv 0$, $\forall k \in \mathbf{N}_{T_f}$ implies $x_0 = 0$.

Definition 3.3. *[Zhang et al. (2008)] The linear discrete-time system*

$$x(k+1) = Ax(k) + Bu(k), \ \ x(0) = x_0, \tag{3.46}$$

$\forall k = k \in \mathbf{N}$ is said to be mean-square stable if for any $x_0 \in \mathbb{R}^n$, the corresponding state satisfies $\lim\limits_{k\to\infty} \|x(k)\| = 0$. The system (3.46) *is said to be stabilizable in the mean square sense if for a constant matrix K, there exists a feedback control law $u(k) = Kx(k)$, that stabilizes the system* (3.46) *mean square stable.*

By using Lemma 3.1, we have the following result under the infinite horizon case as the extension: Suppose that the linear discrete-time system

$$x(k+1) = Ax(k) + Bu(k), \ \ x(0) = x_0, \tag{3.47}$$

$\forall k = k \in \mathbf{N}_{T_f}$ is mean-square stable, where A and B are assumed to be constant matrices of suitable dimensions. Let us define the cost functional

$$J(x_0, \ u) := \sum_{k=0}^{\infty} [x^T(k)Qx(k) + 2x^T(k)Su(k) + u^T(k)Ru(k)], \tag{3.48}$$

where $Q = Q^T \geq 0, \ R = R^T > 0$, and $Q - SR^{-1}S^T > 0$.

Lemma 3.3. *For the discrete-time optimal control problem* (3.47) *with cost functional* (3.48)*, the optimal feedback strategy is given by*

$$u^*(k) = Kx(k) = -\hat{R}^{-1}\big(B^T XA_S + R^{-1}S\big)x(k), \tag{3.49}$$

where $X^T = X \geq 0$ is the solution of the following algebraic Riccati equation (ARE):

$$X = A_S^T XA_S - A_S^T XVXA_S + Q - S^T R^{-1} S, \tag{3.50}$$

with

$$\begin{aligned} &A_S := A - BR^{-1}S, \\ &V := B\hat{R}^{-1}B^T, \\ &\hat{R} := R + B^T XB, \\ &Q - S^T R^{-1} S \geq 0. \end{aligned}$$

Proof. Using optimal control $u^*(k) = Kx(k)$, the state feedback system (3.47) can be written as

$$x(k+1) = (A+BK)x(k), \ \ x(0) = x_0, \tag{3.51}$$

with cost functional

$$J(x_0,\ u^*) := \sum_{k=0}^{\infty}[x^T(k)(Q+2SK+K^TRK)x(k)]. \tag{3.52}$$

Suppose that there exits a symmetric positive definite matrix X such that the ARE (3.56) holds for all admissible control inputs. Let us define the Lyapunov candidate function

$$V(x(k)) = x^T(k)Xx(k), \tag{3.53}$$

where $V(x(k) > 0$ for all $x(k) \neq 0$.

The difference between corresponding trajectory of the system (3.51) is given by

$$\begin{aligned}\Delta V(x(k)) &= V(x(k+1)) - V(x(k)) \\ &= x^T(k+1)Xx(k+1) - x^T(k)Xx(k) \\ &= x^T(k)(A+BK)^TX(A+BK)x(k) - x^T(k)Xx(k) \\ &= x^T(k)[(A+BK)^TX(A+BK) - X]x(k),\end{aligned} \tag{3.54}$$

which is stable if $\Delta V(x(k)) < 0$. Then, we can form the discrete-time Lyapunov stabilizable equation as follows:

$$(A+BK)^TX(A+BK) - X = -(Q+2SK+K^TRK) \tag{3.55}$$

Substituting the value of $K = -\hat{R}^{-1}\left(B^TXA_S + R^{-1}S\right)$ to equation (3.55) and simplifying, we can get the following ARE:

$$X = A_S^TXA_S - A_S^TXVXA_S + Q - S^TR^{-1}S, \tag{3.56}$$

with

$$\begin{aligned}&A_S := A - BR^{-1}S, \\ &V := B\hat{R}^{-1}B^T, \\ &\hat{R} := R + B^TXB, \\ &Q - S^TR^{-1}S \geq 0.\end{aligned}$$

Hence, Lemma 3.3 is proved. □

On the other hand, consider the following discrete-time system.

$$x(k+1) = Ax(k) + Dv(k), \tag{3.57a}$$

$$z(k) = Cx(k), \quad x(0) = x_0, \quad k = 0,\ 1,\ \dots T_f, \tag{3.57b}$$

with performance

$$J_\gamma(x_0,\ v) := \sum_{k=0}^{\infty}[\gamma^2\|v(k)\|^2 - \|z(k)\|^2], \tag{3.58}$$

where $v(k) \in \mathbb{R}^{n_v}$ represents the external disturbance. $z(k) \in \mathbb{R}^{n_z}$ represents the controlled output.

Lemma 3.4. *For the discrete time system* (3.57)*, $\|L_{T_f}\| < \gamma$ for given disturbance attenuation level $\gamma > 0$, the worst-case disturbance is given by*

$$v^*(k) = F_\gamma x(k) = -T_\gamma^{-1}D^T YAx(k), \tag{3.59}$$

if and only if there exists a unique solution $Y \leq 0$ to the following matrix difference equation:

$$Y = A^T YA - A^T YUYA - C^T C, \tag{3.60}$$

where

$$U := DT_\gamma^{-1}D^T,$$
$$T_\gamma := \gamma^2 I_{n_v} + D^T YD.$$

Proof. Using the worst-case disturbance $v^*(k) = F_\gamma x(k)$, the state feedback system (3.57a) can be written as

$$x(k+1) = (A + DF_\gamma)x(k),\ \ x(0) = x_0, \tag{3.61}$$

with cost functional

$$\begin{aligned} J_\gamma(x_0,\ v) &= \sum_{k=0}^{\infty}[\gamma^2 x^T(k)F_\gamma^T F_\gamma x(k) - x^T(k)C^T Cx(k)] \\ &= \sum_{k=0}^{\infty} x^T(k)(\gamma^2 F_\gamma^T F_\gamma - C^T C)x(k). \end{aligned} \tag{3.62}$$

Suppose that there exits a symmetric positive definite matrix Y such that the ARE (3.60) holds for all admissible control inputs. Let us define the Lyapunov candidate function

$$V(x(k)) = x^T(k)Yx(k), \tag{3.63}$$

where $V(x(k) > 0$ for all $x(k) \neq 0$.

The difference between corresponding trajectory of the system (3.57) is given by

$$\Delta V(x(k)) = V(x(k+1)) - V(x(k))$$

$$
\begin{aligned}
&= x^T(k+1)Yx(k+1) - x^T(k)Yx(k) \\
&= x^T(k)(A+DF_\gamma)^T Y(A+DF_\gamma)x(k) - x^T(k)Yx(k) \\
&= x^T(k)[(A+DF_\gamma)^T Y(A+DF_\gamma) - Y]x(k),
\end{aligned} \tag{3.64}
$$

which is stable if $\Delta V(x(k)) < 0$. Then, we can form the discrete-time Lyapunov stabilizable equation as follows:

$$
(A+DF_\gamma)^T Y(A+DF_\gamma) - Y = -(\gamma^2 F_\gamma^T F_\gamma - C^T C) \tag{3.65}
$$

Substituting the value of $F_\gamma = -T_\gamma^{-1}D^T YA$ to equation (3.65) and simplifying, we can get the following ARE:

$$
Y = A^T YA - A^T YUYA - C^T C, \tag{3.66}
$$

where

$$
\begin{aligned}
&U := DT_\gamma^{-1}D^T, \\
&T_\gamma := \gamma^2 I_{n_v} + D^T YD.
\end{aligned}
$$

Hence, Lemma 3.4 is proved. □

Consider a time-invariant linear discrete-time system with multiple follower is described by

$$
x(k+1) = Ax(k) + \sum_{j=1}^{N} B_{0j}u_{0j}(k) + \sum_{j=1}^{N} B_j u_j(k) + Dv(k), \tag{3.67a}
$$

$$
z(k) = \begin{bmatrix} Cx(k) \\ G_0\boldsymbol{u}_0(k) \\ G_1 u_1(k) \\ \vdots \\ G_N u_N(k) \end{bmatrix}, \tag{3.67b}
$$

where $x(0) = x_0,\ G_i^T G_i = I_{m_i}$.
The cost functionals are defined as

$$
\begin{aligned}
&J_0(x_0, \boldsymbol{u}_0,\ u_1,\ \dots\ , u_N,\ v) \\
&\quad := \frac{1}{2}\sum_{k=0}^{\infty}\left[x^T(k)Q_0x(k) + \sum_{j=1}^{N} u_{0j}^T(k)R_{00j}u_{0j}(k) + \sum_{j=1}^{N} u_j^T(k)R_{0j}u_j(k)\right],
\end{aligned} \tag{3.68a}
$$

$$
J_\gamma(x_0, \boldsymbol{u}_0,\ u_1,\ \dots\ , u_N,\ v) := \sum_{k=0}^{\infty}[\gamma^2\|v(k)\|^2 - \|z(k)\|^2], \tag{3.68b}
$$

$$J_i(x_0, \boldsymbol{u}_0, u_1, \dots, u_N, v)$$
$$:= \frac{1}{2}\sum_{k=0}^{\infty}\left[x^T(k)Q_i x(k) + u_{0i}^T(k)R_{0ii}u_{0i}(k) + u_i^T(k)R_{ii}u_i(k)\right],\ i = 1, \dots, N, \quad (3.68c)$$

where $Q_i = Q_i^T \geq 0$, $i = 0, 1, \dots, N$, $R_{00i} = R_{00i}^T > 0$, $R_{0i} = R_{0i}^T \geq 0$, $R_{0ii} = R_{0ii}^T \geq 0$, $R_{ii} = R_{ii}^T > 0$.

In order to solve this problem, first we find the leader's team optimal solution $u_c^*(k)$ attenuating the disturbance under H_∞ constraint. We can accomplish this task by considering $u_c(k)$ and $v(k)$ to perform Nash equilibrium. While considering worst-case disturbance $v^*(k) = F_\gamma x(k)$, we can solve it by Lemma 3.3 for the system (3.67) with cost functional (3.68a); and while considering team-optimal state feedback control $u_c^*(k) = K_c x(k)$ for the system (3.67) with cost functional (3.68b), we can obtain the solution by Lemma 3.4. Hence, the following cross-coupled algebraic Riccati equations (CCAREs) can be found:

$$P = A_v^T P A_v - A_v^T P S_c P A_v + Q_0, \quad (3.69a)$$
$$W = A_K^T W A_K - A_K^T W U_W W A_K - L_K, \quad (3.69b)$$

where

$$\begin{aligned}
A_v &:= A + DF_\gamma,\\
A_K &:= A + B_c K_c,\\
S_c &:= B_c \hat{R}_c^{-1} B_c^T,\\
U_W &:= D T_{W\gamma}^{-1} D^T,\\
F_\gamma &:= -T_{W\gamma}^{-1} D^T W A_K,\\
T_{W\gamma} &:= \gamma^2 I_{n_v} + D^T W D,\\
\hat{R}_c &:= R_c + B_c^T P B_c,\\
L_K &:= C^T C + K_c^T K_c,\\
K_c &:= \begin{bmatrix} \boldsymbol{K}_0 \\ \boldsymbol{K}_1 \end{bmatrix} = -\hat{R}_c^{-1} B_c^T P A_v,\\
\boldsymbol{K}_0 &:= \begin{bmatrix} K_{01}^T & \cdots & K_{0N}^T \end{bmatrix}^T,\\
\boldsymbol{K}_1 &:= \begin{bmatrix} K_1^T & \cdots & K_N^T \end{bmatrix}^T,\\
B_c &:= \begin{bmatrix} \boldsymbol{B}_0 & B_1 & \cdots & B_N \end{bmatrix},\\
R_c &:= \textbf{block diag}\begin{pmatrix} \boldsymbol{R}_0 & R_{01} & \cdots R_{0N} \end{pmatrix}.
\end{aligned}$$

On the other hand, to ensure each i-th follower's optimal state feedback Nash equilibrium strategy for the system (3.67) with cost functional (3.68c), we can use the BDRE (3.42) and the followers' optimal strategy (3.44) determined for the finite-horizon case. According to Lemma 3.3, these results can be extended infinitely and the following CCARE can be established:

$$P_i = \tilde{A}_i^T P_i \tilde{A}_i - \tilde{X}_i^T \tilde{Y}_i^{-1} \tilde{X}_i + \tilde{Q}_i, \tag{3.70a}$$

$$\tilde{K}_i = -[\tilde{Y}_i]^{-1} \tilde{X}_i, \quad i = 1,\ 2,\ \ldots\ ,N, \tag{3.70b}$$

where

$$\begin{aligned}
\tilde{A}_i &:= A + DF_\gamma + B_{0i}\eta_{0i} + \sum_{j=1,\ j\neq i}^{N} [B_{0j}K_{0j} + B_jK_j],\\
\tilde{X}_i &:= \tilde{B}_i^T P_i \tilde{A}_i + \eta_{ii}^T R_{0ii}\eta_{0i},\\
\tilde{Y}_i &:= \tilde{R}_i + \tilde{B}_i^T P_i \tilde{B}_i,\\
\tilde{B}_i &:= B_i + B_{0i}\eta_{ii},\\
\tilde{Q}_i &:= Q_i + \eta_{0i}^T R_{0ii}\eta_{0i},\ \tilde{R}_i := R_{ii} + \eta_{ii}^T R_{0ii}\eta_{ii}.
\end{aligned}$$

Remark 3.2. *It should be noted that the incentive parameter η_{ii} can be uniquely determined if and only if $\tilde{Y}_i$ is non-singular.*

Proposition 3.1. *If there exists a solutions set of the CCAREs* (3.69) *and* (3.70) *then the following strategy-sets for leader, followers and under the worst-case disturbance are defined correspondingly for the two-level incentive Stackelberg game with H_∞ constraint as:*

$$u_{0i}^*(k) := \eta_{0i}x(k) + \eta_{ii}^* u_i^*(k) = K_{0i}x(k), \tag{3.71a}$$

$$u_i^*(k) := \tilde{K}_i x(k), \tag{3.71b}$$

$$v^*(k) := F_\gamma x(k). \tag{3.71c}$$

In order to solve the CCAREs of (3.69) and (3.70), first the following computational algorithm is based on the Lyapunov iteration:

$$\begin{cases} P^{(r+1)} = [A_v^{(r)}]^T P^{(r+1)} A_v^{(r)} - [A_v^{(r)}]^T P^{(r)} S_c^{(r)} P^{(r)} A_v^{(r)} + Q_0, \\ W^{(r+1)} = [A_K^{(r)}]^T W^{(r+1)} A_K^{(r)} - [A_K^{(r)}]^T W^{(r)} U_W^{(r)} W^{(r)} A_K^{(r)} - L_K^{(r)},\ r = 0,\ 1,\ \ldots \end{cases} \tag{3.72a}$$

$$\begin{cases} P_i^{(s+1)} = [\tilde{A}_i^{(s)}]^T P_i^{(s+1)} \tilde{A}_i^{(s)} - [\tilde{X}_i^{(s)}]^T [\tilde{Y}_i^{(s)}]^{-1} \tilde{X}_i^{(s)} + \tilde{Q}_i^{(s)}, \\ [\eta_{ii}^{(s+1)}]^T = -\Big(\big(R_{ii} + B_i^T P_i^{(s+1)} B_{0i}\eta_{ii}^{(s)} + B_i^T P_i^{(s+1)} B_i\big)\tilde{K}_i^{(s)} + B_i^T P_i^{(s+1)} \tilde{A}_i^{(s)} \Big) \\ \times \Big(B_{0i}^T P_i^{(s+1)} B_{0i}\eta_{ii}^{(s)} \tilde{K}_i^{(s)} + B_{0i}^T P_i^{(s+1)} B_i \tilde{K}_i^{(s)} + B_{0i}^T P_i^{(s+1)} \tilde{A}_i^{(s)} + R_{0ii}K_{0i} \Big)^{-1}, \\ s = 0,\ 1,\ \ldots i = 1,\ 2,\ \ldots\ ,N, \end{cases} \tag{3.72b}$$

where $P^{(0)} = P_i^{(0)} = I_n,\ W^{(0)} = -I_n,\ \eta_{ii}^{(0)} = \eta_{(ii)}^{(0)}$,

$$
\begin{aligned}
&A_v^{(r)} := A + DF_\gamma^{(r)},\\
&A_K^{(r)} := A + B_c K_c^{(r)},\\
&S_c^{(r)} := B_c[\hat{R}_c^{(r)}]^{-1} B_c^T,\\
&U_W^{(r)} := D[T_{W\gamma}^{(r)}]^{-1} D^T,\\
&F_\gamma^{(r)} := -[T_{W\gamma}^{(r)}]^{-1} D^T W^{(r)} A_K,\\
&T_{W\gamma}^{(r)} := \gamma^2 I_{n_v} + D^T W^{(r)} D,\\
&\hat{R}_c^{(r)} := R_c + B_c^T P^{(r)} B_c,\\
&L_K^{(r)} := C^T C + [K_c^{(r)}]^T K_c^{(r)},\\
&K_c^{(r)} := -[\hat{R}_c^{(r)}]^{-1} B_c^T P^{(r)} A_v^{(r)},\\
&\tilde{A}_i^{(s)} := A + DF_\gamma + B_{0i}\eta_{0i}^{(s)} + \sum_{j=1,\ j\neq i}^{N} [B_{0j}K_{0j} + B_j K_j],\\
&\tilde{X}_i^{(s)} := \tilde{B}_i^T P_i^{(s)} \tilde{A}_i^{(s)} + [\eta_{ii}^{(s)}]^T R_{0ii}\eta_{0i}^{(s)},\\
&\tilde{Y}_i^{(s)} := \tilde{R}_i + \tilde{B}_i^T P_i^{(s)} \tilde{B}_i,\\
&\tilde{K}_i^{(s)} := -[\tilde{Y}_i^{(s)}]^{-1} \tilde{X}_i^{(s)},\\
&\tilde{B}_i^{(s)} := B_i + B_{0i}\eta_{ii}^{(s)},\ \tilde{Q}_i^{(s)} := Q_i + [\eta_{0i}^{(s)}]^T R_{0ii}\eta_{0i}^{(s)},\\
&\tilde{R}_i^{(s)} := R_{ii} + [\eta_{ii}^{(s)}]^T R_{0ii}\eta_{ii}^{(s)}.
\end{aligned}
$$

It should be noted that the initial guess of η_{ii} has to be chosen appropriately. It should be also noted that the convergence of the algorithm (3.72) is not unclear for the reader. In the next section, a numerical example will show that this algorithm can be worked well in practice.

3.5 Numerical example

In order to demonstrate the efficiency of our proposed three strategies, a simple numerical example is investigated. Here we present the example for infinite-horizon case with two non-cooperative players. Let us consider the following system matrices:

$$
\begin{aligned}
&A = \begin{bmatrix} 0.52 & 1.12 \\ 0 & -0.24 \end{bmatrix},\ B_{01} = \begin{bmatrix} 0.138 & 0.20 \\ -0.55 & 0.84 \end{bmatrix},\ B_{02} = \begin{bmatrix} 0.312 & 1.20 \\ -1.25 & 1.03 \end{bmatrix},\\
&B_1 = \begin{bmatrix} 0.15 & -0.11 \\ 0.12 & 2.28 \end{bmatrix},\ B_2 = \begin{bmatrix} 0.23 & -0.45 \\ -0.52 & 1.02 \end{bmatrix},\\
&D = \begin{bmatrix} 0.054 & -0.076 \\ -0.035 & -0.094 \end{bmatrix},\ C = \begin{bmatrix} 1 & 2 \end{bmatrix},
\end{aligned}
$$

$$Q_0 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},\ Q_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},\ Q_2 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix},$$

$$R_{001} = 1.9I_2,\ R_{002} = 2.5I_2,\ R_{01} = 2.7I_2,\ R_{02} = 3.5I_2,$$

$$R_{011} = 4.8I_2,\ R_{022} = 5I_2,\ R_{11} = 0.3I_2,\ R_{22} = 0.5I_2.$$

We choose the disturbance attenuation level as $\gamma = 5$. First, the CCAREs (3.69a) and (3.69b) are solved by using the algorithm (3.72a). These solutions that attain the H_∞-constrained team-optimal solutions are given below:

$$P = \begin{bmatrix} 1.1667 & 0.3607 \\ 0.3607 & 2.7939 \end{bmatrix},\ W = \begin{bmatrix} -1.0756 & -2.1544 \\ -2.1544 & -4.3203 \end{bmatrix},$$

$$K_c = \begin{bmatrix} -0.0252 & -0.0708 \\ -0.0308 & -0.0418 \\ -0.0433 & -0.1219 \\ -0.1511 & -0.3038 \\ -0.0175 & -0.0354 \\ 0.0186 & 0.0879 \\ -0.0201 & -0.0351 \\ 0.0431 & 0.1098 \end{bmatrix},\ F_\gamma = \begin{bmatrix} -0.0001 & -0.0003 \\ -0.0024 & -0.0047 \end{bmatrix}.$$

Second, the CCAREs (3.70a) and (3.70b) are solved by using the algorithm (3.72b).

$$P_1 = \begin{bmatrix} 2.1814 & 0.3941 \\ 0.3941 & 1.8618 \end{bmatrix},\ P_2 = \begin{bmatrix} 1.2091 & 0.9408 \\ 0.9408 & 3.9370 \end{bmatrix},$$

$$\eta_{11} = \begin{bmatrix} 1.6976 & 0.6457 \\ 1.1122 & 1.1856 \end{bmatrix},\ \eta_{22} = \begin{bmatrix} -0.3909 & -0.9033 \\ 0.8427 & 0.8926 \end{bmatrix}.$$

The algorithm (3.72b) converges to the required solution with an accuracy of $1.0e-12$ order after ten iterations. It should be noted that the incentive strategy (3.71a) that will be announced by the leader can calculated as

$$u_{0i}(k) = \eta_{0i}^* x(k) + \eta_{ii}^* u_i(k), \tag{3.73}$$

where

$$\eta_{01} = \begin{bmatrix} -0.0075 & -0.0675 \\ -0.0334 & -0.1066 \end{bmatrix},\ \eta_{02} = \begin{bmatrix} -0.0122 & -0.0364 \\ -0.1727 & -0.3722 \end{bmatrix}.$$

In fact, after announcing this incentive, the followers' strategy can be computed by applying the standard LQ theory

$$u_i^*(k) = [\tilde{R}_i + \tilde{B}_i^T P_i \tilde{B}_i]^{-1} [\tilde{B}_i^T P_i \tilde{A}_i + \eta_{ii}^T R_{0ii} \eta_{0i}] x(k), \tag{3.74}$$

which implies

$$u_1^*(k) = \begin{bmatrix} -0.0175 & -0.0354 \\ 0.0186 & 0.0879 \end{bmatrix} x(k),$$

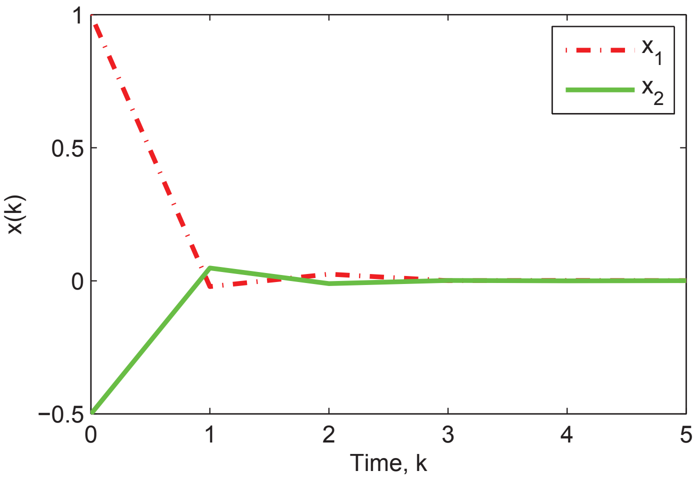


Fig. 3.1: Trajectory of the state.

$$u_2^*(k) = \begin{bmatrix} -0.0201 & -0.0351 \\ 0.0431 & 0.1098 \end{bmatrix} x(k).$$

Indeed, it can be observed that this matrix gain is equal to $\tilde{K}_i$. Namely, it can be confirmed that the followers take the team-optimal solution with H_∞ constraint eventually. In fact, after announcing this incentive, the followers' strategy can be computed by applying the standard LQ theory.

$$u_i^*(k) = -[\tilde{R}_i + \tilde{B}_i^T P_i \tilde{B}_i]^{-1} [\tilde{B}_i^T P_i \tilde{A}_i + \eta_{ii}^T R_{0ii} \eta_{0i}] x(k), \tag{3.75}$$

Fig. 3.1 shows the response of the system with a state trajectory. In addition, Fig. 3.1 represents that the state variables $x(k)$ can stabilize the given system, which implies that the proposed method is very useful and reliable.

3.6 Conclusion

This chapter investigates the incentive Stackelberg game for discrete-time deterministic systems. However, stochastic systems are not considered here. It is the motivation to choose a deterministic system to extend it to a stochastic system. This chapter studies the most common linear quadratic (LQ) optimal control in the game problems. In order to solve the LQ problem, discrete-time maximum principle is deeply studied. This chapter involves one leader and multiple followers rewarding the Starkberg game. For this game, incentive Stackelberg strategy is a broad idea, and leaders can implement his team-optimal solution in a Stackelberg game. In the followers' group, players are supposed to be non-cooperative; subsequently, Nash equilibrium is investigated.

The deterministic disturbances and their attenuation to systems under the H_∞ constraint is the main attraction of this chapter. Problems involving deterministic disturbance must be attenuated at a given target called disturbance attenuation level $\gamma > 0$. Surprisingly, the concept of solving the disturbance reduction problem under the H_∞ constraint seems like a Nash equilibrium between the disturbance input and the control input. In this game, an incentive structure is developed in such a way that leader achieve team-optimal solution attenuating the disturbance under H_∞ constraint. Simultaneously, followers achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance is considered.

This chapter also derives results based on a structure similar to the finite time domain case under infinite time domain conditions. In an infinite-horizon case, incentive Stackelberg game with one leader and multiple followers has also investigated for a discrete time systems with H_∞ constraint. Leader's team-optimal solution attenuating the disturbance under H_∞ constraint is also implemented. On the other hand, followers ensure their Nash equilibrium under the leader's incentives considering the worst-case disturbance. The main attraction of the infinite horizon situation is Lyapunov stability theory. Using Lyapunov stability theory, several theorems and lemmas have been proved.

In this chapter, the team-optimal solution for the leader is achieved in contrast to multiple non-cooperative followers' optimal state feedback gain. The sufficient condition for optimality according to the followers' act subject to the Nash equilibrium condition was also verified. The solution sets for incentive Stackelberg strategy are found by solving a set of backward difference Riccati equations (BDREs) in the finite-horizon case. On the other hand, it is shown that the results of the infinite-horizon case are found by solving a set of algebraic Riccati equations (AREs). An algorithm based on Lyapunov iterations is developed to obtain a solution set of the coupled algebraic Riccati equations. In order to ensure the stability of the system, the state trajectory figure is presented. To demonstrate the effectiveness of the proposed method, a numerical example is demonstrated. However, this chapter only investigates one leader, which leads many leaders to further study.

Chapter 4

H_∞-Constrained Incentive Stackelberg Games for Discrete-Time Stochastic Systems with Multiple Followers

This chapter is based on a previously published article [Ahmed et al. (2017a)].

4.1 Introduction

The Stackelberg game is a strategic game in which a leader declare his/her strategy first. Then, followers perform their optimal decisions subject to the leader's announcement. Finally, the leader will modifies his/her action confirming the followers' response. Subsequently, this two-player static game was extended to a dynamic game with different information patterns [Chen and Cruz (1972), Simaan et al. (1973)]. Among the information patterns, closed-loop Stackelberg strategies with applications were attracting considerable research interest as - *linear quadratic* (LQ) problems [Medanic (1978), Basar and Selbuz (1979), Tolwinski (1981)]. The idea of team-optimal solutions opens new directions for closed-loop Stackelberg strategies. In [Basar and Olsder (1980)], necessary and sufficient conditions for both finite- and infinite-horizon closed-loop feedback solutions were derived for a team problem in which all players optimized a leader's cost functional jointly. Furthermore, [Salman and Cruz (1983)] derived team-optimal closed-loop Stackelberg strategies for systems with slow and fast modes.

The purpose of the incentive mechanism is to induce virtual cooperation in non-cooperative followers so that optimal system performance (reflected in the leader's objective function) is achieved through hierarchical decision-making [Saksena and Cruz (1985)]. An incentive Stackelberg strategy is one where the leader achieves their team-optimal solution to the hierarchical game by using an incentive mechanism. The following two steps are the main elements of an incentive Stackelberg problem [Ho et al. (1982), Basar and Olsder (1980)]. i) The leader determines a team-optimal strategy-set and announces it ahead of time. ii) Knowing the incentive, based on the leader's announced team-optimal strategy, each follower chooses a strategy so as to minimize their own cost. It should be noted that no matter how the followers behave, the leader can achieve their own team-optimal equilibrium by using the corresponding incentive strategy-set. Incentive Stackelberg games apply to organizations with several participants and with organizational objective functions that may not be the same as the members' objective functions. In the theory of teams, each member of the organization has access to different information. In this game, by contrast, it is an important feature that the leader is able to induce followers to cooperate with them as a team, with the leader's objective function as the objective function of the team, while the followers also optimize their own objective functions [Salman and Cruz (1981)]. Incentive Stackelberg strategies have been extensively studied for more than 30 years (see e.g. [Salman and Cruz (1981), Ho et al. (1982), Saksena and Cruz (1985), Zheng and Basar (1982), Zheng et al. (1984)] and references therein). In [Mizukami and Wu (1988)], incentive Stackelberg games with one leader and two non-cooperative followers were solved for an LQ differential game. In [Li et al. (2002)], a team-optimal state feedback incentive Stackelberg strategy for discrete-time two-player nonzero-sum dynamic games was developed for LQ problems. However, none of those studies have considered stochastic noise and deterministic disturbances in the system, which make the problem more challenging.

In recent years, incentive Stackelberg games with robust control theory have been studied for discrete-time linear systems in [Ahmed and Mukaidani (2016), Mukaidani et al. (2017c)]. In [Ahmed and Mukaidani (2016)], one leader and multiple non-cooperative followers are considered a deterministic system, whereas our current study focuses stochastic systems. Unlike [Mukaidani et al. (2017c)], where one leader and one follower are considered a stochastic system, this Chapter deals with one leader and multiple non-cooperative followers. Similar to [Ahmed and Mukaidani (2016)] and [Mukaidani et al. (2017c)], a deterministic disturbance is considered in this Chapter, which is also seen in [Mukaidani et al. (2017d)]. On the other hand, continuous-time stochastic systems are investigated for an infinite-horizon incentive Stackelberg game in [Mukaidani (2016)], where multiple non-cooperative leaders are considered. In [Mukaidani and Xu (2018)], an incentive Stackelberg strategy for continuous-time stochastic linear systems with exogenous disturbances is derived. One leader and multiple non-cooperative followers are considered there and no discussion is included for a discrete-time case yet. This is one of the vital reasons that motivates us to investigate our current study. Accordingly, this Chapter might be viewed as a discrete-time version of [Mukaidani and Xu (2018)].

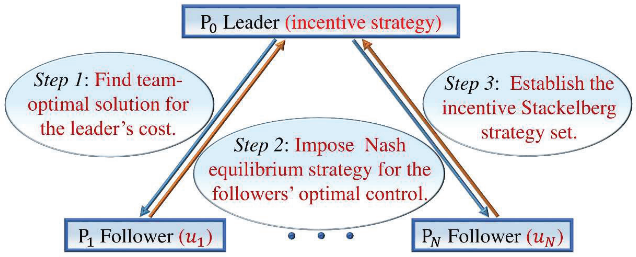


Fig. 4.1: Structure of the incentive Stackelberg game.

This Chapter investigates incentive Stackelberg games with one leader and multiple non-cooperative followers for a discrete-time stochastic system with an exogenous deterministic disturbance, which will be attenuated under the H_∞ constraint. We discuss only two-level hierarchical games with one leader and multiple non-cooperative followers and the hierarchical structure is depicted in Fig. 4.1. Among multiple players P_i, $i = 0, 1, \ldots, N$; P_0 is considered as the leader and $P_1, \ldots, P_N$ are considered as the followers, under the specification that each follower acts non-cooperatively.

We prove that the leader's discrete-time incentive Stackelberg strategies exist under an H_∞ constraint, based on existing results for finite-horizon H_2/H_∞ discrete-time stochastic systems [Zhang et al. (2007)], by means of the state feedback information structure. It should be noted that this information structure seems to be conservative. However, as an engineering application of incentive Stackelberg strategies, a scheduling problem involving a packet switch operating in a ring architecture has been introduced [Saksena and Cruz (1985)]. In this problem, the leader represents the central processor and the followers represent the local link controllers. The information structure of the problem is such that the leader has access to both the decision values and observations of all the followers at each stage of the process. The design of incentive mechanisms that induce non-cooperative followers to virtually cooperate in achieving some system-wide goal is an important feature of hierarchical decision-making. Therefore, a state feedback information structure is sufficient to guarantee the existence of an incentive strategy. It is shown that a solution can be found by solving four cross-coupled *stochastic matrix-valued difference equations* (SMVDEs) and a *stochastic back-ward difference Riccati equation* (SBDRE) in the finite-horizon case. Moreover, the Nash equilibrium strategies of the followers are derived in such a way that ensure the leader's team-optimal solution. Apart from the finite-horizon case, four cross-coupled *stochastic matrix-valued algebraic equations*

(SMVAEs) corresponding to the existing result in [Zhang et al. (2008)] and a *stochastic algebraic Riccati equation* (SARE) are derived in the infinite horizon case to determine the leader's discrete-time incentive Stackelberg strategies. A recursive algorithm based on the Lyapunov iteration is also designed to ease the complexity of computation. Finally, an academic and a practical numerical examples demonstrate the efficiency of the proposed methodology.

This is the first attempt to formulate a one leader/multiple followers discrete-time linear stochastic control problem with an external disturbance. Our main contributions, demonstrated in the subsequent discussion, are as follows. i) The addition of disturbances to this type of problem, and their attenuation under an H_∞ constraint, is considered for the first time. ii) Multiple non-cooperative followers in such a discrete-time stochastic system are considered for the first time. iii) We realize a team-optimal solution for the leader, while simultaneously guaranteeing the Nash equilibrium states of the followers. iv) To solve this problem, a recursive algorithm based on the Lyapunov equation has been developed. v) The system model is new and quite comprehensive in that it covers time-varying parameters, stochastic control schemes, state-multiplicative noise, and exogenous disturbance inputs, thereby more closely reflecting real-world systems.

Notation: The notations used in this Chapter are fairly standard. I_n denotes the $n \times n$ identity matrix; **block diag**$(\cdot)$ denotes the block diagonal matrix; **diag**$(\cdot)$ denotes the diagonal matrix; δ_{ij} denotes the Kronecker delta; $\mathbb{E}[\cdot]$ denotes the expectation operator, $\mathbf{N}_{T_f} := \{0,\ 1,\ \ldots,T_f\}$ and $\mathbf{N} := \{0,\ 1,\ \ldots\}$. The l^2-norm of $y(\cdot) \in l^2_w(\mathbf{N}_{T_f}, \mathbb{R}^n)$ is defined by

$$\|y(\cdot)\|^2_{l^2_w(\mathbf{N}_{T_f},\mathbb{R}^n)} := \sum_{k=0}^{T_f} \mathbb{E}[\|y(k)\|^2].$$

Finally, define an N-tuple

$$\gamma := (\gamma_1,\ \ldots,\gamma_N) \in \Gamma_1 \times \ldots \times \Gamma_N,$$

for given sets Γ_i, we write

$$\gamma^*_{-i}(\alpha) := (\gamma^*_1,\ \ldots,\gamma^*_{i-1},\ \alpha,\ \gamma^*_{i+1},\ \ldots,\gamma^*_N),$$

where the superscript $(^*)$ is used in the optimal case.

4.2 Definitions and preliminaries

Consider a discrete-time stochastic system with the deterministic disturbance.

$$\begin{cases} x(k+1) = A(k)x(k) + B(k)u(k) + D(k)v(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} C(k)x(k) \\ G(k)u(k) \end{bmatrix}, \; G^T(k)G(k) = I_{n_u}, \; k \in \mathbf{N}_{T_f}, \end{cases} \tag{4.1}$$

where $u(k) \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_u})$ represents the control input, $v(k) \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_v})$ represents the disturbance input and $z(k) \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_z})$ represents the controlled output, $w(k)$ is a real-valued random variable defined in the filtered probability space, second-order process with $\mathbb{E}[w(k)] = 0$ and $\mathbb{E}[w(s)w(k)] = \delta_{sk}$ [Zhang et al. (2007), Zhang et al. (2008)].

Given the disturbance attenuation level $\gamma > 0, 0 < T_f < \infty$, an optimal state feedback control $u^*(k) \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_u})$ defined as

$$u^*(k) := K(k)x(k) \in \mathbb{R}^{n_u},$$

can be found (if existing) such that:
(i) For the closed-loop system

$$\begin{cases} x(k+1) = (A(k) + B(k)K(k))x(k) + D(k)v(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} C(k)x(k) \\ G(k)K(k)x(k) \end{bmatrix}, \; G^T(k)G(k) = I_{n_u}, \; k \in \mathbf{N}_{T_f}, \end{cases} \tag{4.2}$$

the following condition holds:

$$\|L_{T_f}\|_{H_\infty} := \sup_{\substack{v \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_v}), \\ v \neq 0, \; x_0 = 0}} \frac{\|z\|_{l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_z})}}{\|v\|_{l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_v})}} < \gamma, \tag{4.3}$$

where

$$\|z\|_{l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_z})} := \left(\sum_{k=0}^{T_f} \mathbb{E}\left[x^T(k)C^T(k)C(k)x(k) + x^T(k)K^T(k)K(k)x(k) \right] \right)^{\frac{1}{2}},$$

$$\|v\|_{l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_v})} := \left(\sum_{k=0}^{T_f} \mathbb{E}\left[v^T(k)v(k) \right] \right)^{\frac{1}{2}}.$$

(ii) For the worst-case disturbance $v^*(k) \in l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_v})$, if existing, is implemented in (4.1), $u^*(k)$ optimizes the cost

$$J_u(u, v^*) := \|z\|^2_{l_w^2(\mathbf{N}_{T_f}, \mathbb{R}^{n_z})} = \sum_{k=0}^{T_f} \mathbb{E}\left[x^T(k)C^T(k)C(k)x(k) + x^T(k)K^T(k)K(k)x(k) \right]. \tag{4.4}$$

When such $(u^*,\ v^*)$ exists, we say that the finite horizon H_2/H_∞ control is solvable. Here, the worst-case disturbance means

$$v^*(k) := \arg\min_v J_v(u^*, v), \tag{4.5}$$

where

$$J_v(u,\ v) := \gamma^2 \|v\|^2_{l^2_w(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})} - \|z\|^2_{l^2_w(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})} \tag{4.6}$$

is associated with the system (4.1). Then, the finite horizon H_2/H_∞ control is equivalent to the following Nash equilibrium $(u^*,\ v^*)$:

$$J_u(u^*,\ v^*) \le J_u(u,\ v^*), \tag{4.7a}$$

$$J_v(u^*,\ v^*) \le J_v(u^*,\ v). \tag{4.7b}$$

Lemma 4.1. *[Zhang et al. (2007)] For given disturbance attenuation level $\gamma > 0$, the finite horizon H_2/H_∞ control for the system* (4.1) *has solutions $(u^*,\ v^*)$ as*

$$u^*(k,\ x(k)) := K(k)x(k),$$
$$v^*(k,\ x(k)) := F_\gamma(k)x(k),$$

with $K(k) \in \mathbb{R}^{n_u \times n}$ and $F_\gamma(k) \in \mathbb{R}^{n_v \times n}$, $k \in \mathbf{N}_{T_f}$ being matrix-valued functions, iff the following four cross-coupled SMVDEs have solutions $(P(k),\ W(k);\ K(k),\ F_\gamma(k))$ with $P(k) \ge 0$ and $W(k) \le 0$, $k \in \mathbf{N}_{T_f}$:

$$\begin{aligned} P(k) = {} & A_v^T(k)P(k+1)A_v(k) + A_p^T(k)P(k+1)A_p - A_v^T(k)P(k+1)B(k) \\ & \times \hat{R}^{-1}(k)B^T(k)P(k+1)A_v(k) + C^T(k)C(k), \quad P(T_f+1) = 0, \end{aligned} \tag{4.8a}$$

$$K(k) = -\hat{R}^{-1}(k)B^T(k)P(k+1)A_v(k), \tag{4.8b}$$

$$\begin{aligned} W(k) = {} & A_u^T(k)W(k+1)A_u(k) + A_p^T(k)W(k+1)A_p(k) - A_u^T(k)W(k+1)U(k)W(k+1)A_u(k) \\ & - C^T(k)C(k) - K^T(k)K(k),\ \ W(T_f+1) = 0, \end{aligned} \tag{4.8c}$$

$$F_\gamma(k) = -T_\gamma^{-1}(k)D^T(k)W(k+1)A_u(k), \tag{4.8d}$$

where

$$\begin{aligned} A_v(k) &:= A(k) + D(k)F_\gamma(k), \\ A_u(k) &:= A(k) + B(k)K(k), \\ U(k) &:= D(k)T_\gamma^{-1}(k)D^T(k), \\ \hat{R}(k) &:= I_{n_u} + B^T(k)P(k+1)B(k) > 0, \\ T_\gamma(k) &:= \gamma^2 I_{n_v} + D^T(k)W(k+1)D(k) > 0. \end{aligned}$$

Proof. *Necessary condition* Substituting $u^*(k,\ x(k)) = K(k)x(k)$ in (4.1), we obtain (4.2). By Lemma 3 of [Zhang et al. (2007)], (4.8c) has a unique solution $W(k) \geq 0$. From the sufficient condition of Lemma 3 of [Zhang et al. (2007)], the worst-case disturbance $v^*(k,\ x(k))$ can be determined by

$$v^*(k,\ x(k)) = F_\gamma(k)x(k) = -T_\gamma^{-1}(k)D^T(k)W(k+1)A_u(k)x(k). \tag{4.9}$$

On the other hand, implementing $v(k,\ x(k)) = v^*(k,\ x(k)) = F_\gamma(k)x(k)$ in (4.1), we obtain

$$\begin{cases} x(k+1) = A(k)x(k) + B(k)u(k) + D(k)F_\gamma(k)x(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} C(k)x(k) \\ G(k)u(k) \end{bmatrix},\ G^T(k)G(k) = I_{n_u},\ k \in \mathbf{N}_{T_f}. \end{cases} \tag{4.10}$$

While the problem is to minimize the linear quadratic cost functional $J_u(u,\ v^*)$, it turns to be a standard discrete-time LQ optimal problem. By Theorem 2.1, (4.8a) provides a unique solution $P(k) \geq 0$. Moreover, the optimal state feedback control can be determined by

$$u^*(k,\ x(k)) = K(k)x(k) = -\hat{R}^{-1}(k)B^T(k)P(k+1)A_v(k)x(k). \tag{4.11}$$

Sufficient condition Substituting $u^*(k,\ x(k)) = K(k)x(k)$ in (4.1), we obtain (4.2). From (4.8c) and Lemma 3 of [Zhang et al. (2007)], we have $\|L_{T_f}\|_{H_\infty} < \gamma$. By Lemma 2 of [Zhang et al. (2007)] and (4.8c), we obtain

$$J_v(u^*,\ v) = \sum_{k=0}^{T_f} \mathbb{E}[\gamma^2\|v(k)\|^2 - \|z(k)\|^2] \geq J_v(u^*,\ v^*) = x_0^T W(0)x_0. \tag{4.12}$$

Therefore, form (4.12), we see that $v^*(k,\ x(k)) = F_\gamma(k)x(k)$ is the worse case disturbance. Similarly, it can be shown that

$$J_u(u,\ v^*) = \sum_{k=0}^{T_f} \mathbb{E}[\|z(k)\|^2 \geq J_u(u^*,\ v^*) = x_0^T P(k)x_0. \tag{4.13}$$

Therefore,

$$J_u(u^*,\ v^*) \leq J_u(u,\ v^*), \tag{4.14a}$$

$$J_v(u^*,\ v^*) \leq J_v(u^*,\ v), \tag{4.14b}$$

which imply that the strategy pair $(u^*,\ v^*)$ solves the finite horizon H_2/H_∞ control problem for the system (4.1). □

Lemma 4.2. *Consider the linear discrete-time stochastic system*

$$x(k+1) = A(k)x(k) + B(k)u(k) + A_p(k)x(k)w(k),\ x(0) = x_0, \quad k \in \mathbf{N}_{T_f}, \tag{4.15}$$

where $u(k) \in \mathbb{R}^m$ *denotes the control input.* $A(k)$*,* $B(k)$ *and* $A_p(k)$ *are assumed to be matrix-valued functions of suitable dimensions.*

Let us define the cost functional

$$J(x_0,\ u) := \sum_{k=0}^{T_f} \mathbb{E}[x^T(k)Q(k)x(k) + 2x^T(k)S(k)u(k) + u^T(k)R(k)u(k)], \tag{4.16}$$

where $Q(k) = Q^T(k) \geq 0,\ R(k) = R^T(k) > 0$*,* $Q(k) - S(k)R^{-1}(k)S^T(k) > 0$ *and* $0 < T_f < \infty$*.*

Then there exists a matrix-valued function $P(k) > 0$ *that solves the following SBDRE of the system* (4.15)–(4.16)*:*

$$\begin{aligned} P(k) =& A_s^T(k)P(k+1)A_s(k) + A_p^T(k)P(k+1)A_p(k) - A_s^T P(k+1)B(k) \\ & \times \hat{R}^{-1}(k)B^T(k)P(k+1)A_s(k) + Q_s(k), \quad P(T_f+1) = 0, \end{aligned} \tag{4.17}$$

where

$$\begin{aligned} A_s(k) &:= A(k) - B(k)R^{-1}(k)S^T(k), \\ \hat{R} &:= R(k) + B^T(k)P(k+1)B(k), \\ Q_s(k) &:= Q(k) - S(k)R^{-1}(k)S^T(k), \end{aligned}$$

and the optimal state feedback control in this case is given by

$$u^*(k) = K(k)x(k) = -\hat{R}^{-1}(k)\left[S^T(k) + B^T(k)P(k+1)A(k)\right]x(k). \tag{4.18}$$

Proof. To prove Lemma 4.2, we use the matrix Lagrangian multiplier method as in [Rami and Zhou (2000)]. At first, we transfer the LQ problem (4.15)–(4.16) in terms of the state covariance matrices $X(k) = \mathbb{E}[x^T(k)x(k)]$. For this instance, we use a closed-loop state feedback control law

$$u(k) = K(k)x(k), \qquad \text{for } k = 1,\ ...,\ T_f, \tag{4.19}$$

where $K(k)$ is the gain matrix for any $x_0 \in \mathbb{R}^n$, the closed-loop system

$$x(k+1) = [A(k) + B(k)K(k)]x(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \tag{4.20}$$

with cost functional

$$J(x_0,\ u) := \sum_{k=0}^{T_f} \mathbb{E}[x^T(k)Q(k)x(k) + 2x^T(k)S(k)K(k)x(k) + K^T(k)x^T(k)R(k)K(k)x(k)]. \tag{4.21}$$

Then, by a simple calculation an equivalent form of the state equation (4.15) can be written as

$$X(k+1) = A(k)X(k)A^T(k) + B(k)K(k)X(k)A^T(k) + A(k)X(k)K^T(k)B^T(k) + B(k)K(k)X(k)K^T(k)B^T(k) + A_p(k)X(k)A_p^T(k), \tag{4.22}$$

and the equivalent cost functional of (4.16) is

$$J(X(k)) = \min_{K(0),\ \dots,\ K(N)\in\mathbb{R}^{m\times n}} \mathbf{Tr}[\{Q(k) + 2S(k)K(k) + K^T(k)R(k)K(k)\}X(k)]. \tag{4.23}$$

Here, the system (4.22)–(4.23) seems to be a deterministic optimal control problem. The Lagrangian function can be represented as follows,

$$\mathbb{L} = \sum_{k=0}^{T_f} H_k, \tag{4.24}$$

where

$$H_k = \mathbf{Tr}[\{Q(k) + 2S(k)K(k) + K^T(k)R(k)K(k)\}X(k)] + \mathbf{Tr}[P(k+1)\{A(k)X(k)A^T(k) + B(k)K(k)X(k)A^T(k) + A(k)X(k)K^T(k)B^T(k) + B(k)K(k)X(k)K^T(k)B^T(k) + A_p(k)X(k)A_p^T(k) - X(k+1)\}], \tag{4.25}$$

and the matrices $P(0),\ \dots,\ P(T_f)$ are the Lagrangian multipliers. The first-order necessary conditions for optimality are

$$\frac{\partial H_k}{\partial K(k)} = 0,$$
$$\frac{\partial H_k}{\partial X(k)} = P(k),\ \text{for } k = 1,\ \dots,\ T_f,\ P(T_f+1) = 0.$$

The calculation of the above derivatives leads to the following equations:

$$[R(k) + B^T(k)P(k+1)B(k)]K(k) + S^T(k) + B^T(k)P(k+1)A(k) = 0, \tag{4.26}$$

$$P(k) = Q(k) + A^T(k)P(k+1)A(k) + A_p^T(k)P(k+1)A_p(k) + K^T(k)[R(k) + B^T(k)P(k+1)B(k)]K(k) + K^T(k)[S^T(k) + B^T(k)P(k+1)A(k)] + [S(k) + A^T(k)P(k+1)B(k)]K(k),\ P(T_f+1) = 0. \tag{4.27}$$

Now, by using Lemma 3.1 of [Rami et al. (2002)], we can see that the existence of a solution $K(0),\ \dots,\ K(T_f)$ to equation (4.26) and the solution is given by the following deterministic gain matrices:

$$K(k) = -[R(k) + B^T(k)P(k+1)B(k)]^{-1}[S^T(k) + B^T(k)P(k+1)A(k)]. \tag{4.28}$$

Transferring the gains from equation (4.28) into equation (4.27), we obtain the following backward difference formula for the Lagrangian multipliers:

$$\begin{aligned} P(k) = {} & A^T(k)P(k+1)A(k) + A_p^T(k)P(k+1)A_p(k) - [S(k) + A^T(k)P(k+1)B(k)] \\ & \times [R(k) + B^T(k)P(k+1)B(k)]^{-1}[S^T(k) + B^T(k)P(k+1)A(k)] \\ & + Q(k),\ P(T_f+1) = 0. \end{aligned} \tag{4.29}$$

Alternatively, equation (4.29) can be rewritten as,

$$\begin{aligned} P(k) = & A_s^T(k)P(k+1)A_s(k) + A_p^T(k)P(k+1)A_p(k) \\ & - A_s^T P(k+1)B(k)\hat{R}^{-1}R(k)B^T(k)P(k+1)A_s(k) + Q_s(k), \quad P(T_f+1) = 0, \end{aligned} \tag{4.30}$$

where

$$\begin{aligned} A_s(k) &:= A(k) - B(k)R^{-1}(k)S^T(k), \\ \hat{R}(k) &:= R(k) + B^T(k)P(k+1)B(k), \\ Q_s(k) &:= Q(k) - S^T(k)R^{-1}(k)S(k). \end{aligned}$$

Using equation (4.28), the closed-loop optimal state feedback control can be written as

$$u^*(k) = K(k)x(k) = -\hat{R}^{-1}(k)\left[S^T(k) + B^T(k)P(k+1)A(k)\right]x(k). \tag{4.31}$$

To complete the proof of the lemma, we need to show now $\hat{R}(k) \geq 0,\ k \in \mathbf{N}_{T_f}$. Let us suppose that there exists $\hat{R}(l)$ associated with a negative eigenvalue λ. Denote the unitary eigenvector corresponding to λ as v_λ (i.e., $v_\lambda^T v_\lambda = 1$ and $\hat{R}(l)v_\lambda = \lambda v_\lambda$). Let $\delta \neq 0$ be an arbitrary scalar and construct a control sequence as follows,

$$\hat{u}(k) = \begin{cases} -\hat{R}^{-1}(k)\hat{S}(k)x(k) & k \neq l, \\ \delta|\lambda|^{-\frac{1}{2}}v_\lambda - \hat{R}^{-1}(k)\hat{S}(k)x(k) & k = l, \end{cases} \tag{4.32}$$

where $\hat{S}(k) := S^T(k) + B^T(k)P(k+1)A(k)$. The associated cost functional is

$$\begin{aligned} J(x_0,\ \hat{u}(0),\ \ldots,\ \hat{u}(T_f)) &= \mathbb{E}\left[\sum_{k=0}^{T_f}[\hat{u}(k) + \hat{R}(k)^{-1}\hat{S}(k)x(k)]^T\hat{R}(k)[\hat{u}(k) \right. \\ & \qquad \left. + \hat{R}(k)^{-1}\hat{S}(k)x(k)] + x_0^T P(0)x_0\right] \\ &= \left(\frac{\delta}{|\lambda|^{\frac{1}{2}}}v_\lambda\right)^T \hat{R}(l)\left(\frac{\delta}{|\lambda|^{\frac{1}{2}}}v_\lambda\right) + \mathbb{E}[x_0^T P(0)x_0] \\ &= -\delta^2 + \mathbb{E}[x_0^T P(0)x_0]. \end{aligned}$$

Definitely, as $\delta \to \infty$, $J(x_0,\ \hat{u}(0),\ \ldots,\ \hat{u}(T_f)) \to -\infty$ which contradicts our assumption. □

Definition 4.1. *[Zhang et al. (2008)] The following discrete-time stochastic system:*

$$\begin{cases} x(k+1) = Ax(k) + A_p x(k)w(k), \\ z(k) = Cx(k), \quad x(0) = x_0 \in \mathbb{R}^n, \quad k \in \mathbf{N}, \end{cases} \tag{4.33}$$

or $(A,\ A_p/C)$ is said to be exactly observable if $z(k) \equiv 0$, $\forall k \in \mathbf{N}$ implies $x_0 = 0$.

Definition 4.2. *[Zhang et al. (2008)] The linear discrete-time stochastic system*

$$x(k+1) = Ax(k) + Bu(k) + A_p x(k)w(k),\ \ x(0) = x_0, \tag{4.34}$$

$\forall k = k \in \mathbf{N}$ is said to be mean-square stable if for any $x_0 \in \mathbb{R}^n$, the corresponding state satisfies $\lim\limits_{k\to\infty} \mathbb{E}\|x(k)\| = 0$. The system (4.34) is said to be stabilizable in the mean square sense if there exists a mean-square feedback stabilizing control law $u(k) = Kx(k)$, where K is a constant matrix.

By using Lemma 4.2, we have the following result under the infinite horizon case as the extension:

Suppose that the linear discrete-time stochastic system

$$x(k+1) = Ax(k) + Bu(k) + A_p x(k)w(k),\ \ x(0) = x_0, \tag{4.35}$$

$\forall k = k \in \mathbf{N}$ is mean-square stable, where $A,\ B$ and A_p are assumed to be constant matrices of suitable dimensions.

Let us define the cost functional

$$J(x_0,\ u) := \sum_{k=0}^{\infty} \mathbb{E}[x^T(k)Qx(k) + 2x^T(k)Su(k) + u^T(k)Ru(k)], \tag{4.36}$$

where $Q = Q^T \geq 0,\ R = R^T > 0$, and $Q - SR^{-1}S^T > 0$.

Lemma 4.3. *There exists a symmetric constant matrix $P > 0$ that solves the following SARE of the system* (4.35)–(4.36)*:*

$$P = A_s^T P A_s + A_p^T P A_p - A_s^T P B \hat{R}^{-1} B^T P A_s + Q_s, \tag{4.37}$$

where

$$\begin{aligned} A_s &:= A - BR^{-1}S^T, \\ \hat{R} &:= R + B^T P B, \\ Q_s &:= Q - SR^{-1}S^T, \end{aligned}$$

and the optimal state feedback control in this case is given by

$$u^*(k) = Kx(k) = -\hat{R}^{-1}\left[S^T + B^T P A\right]x(k). \tag{4.38}$$

Proof. Using optimal control $u^*(k) = Kx(k)$, the state feedback system (4.35) can be written as

$$x(k+1) = (A+BK)x(k) + A_p x(k)w(k), \;\; x(0) = x_0, \tag{4.39}$$

with cost functional

$$J(x_0,\, u^*) := \sum_{k=0}^{\infty} \mathbb{E}[x^T(k)(Q+2SK+K^T RK)x(k)]. \tag{4.40}$$

Suppose that there exits a symmetric positive definite matrix P such that the SARE (4.37) holds for all admissible control inputs. Let us define the Lyapunov candidate function

$$\mathbb{E}[V(x(k))] = \mathbb{E}[x^T(k)Px(k)], \tag{4.41}$$

where $V(x(k) > 0$ for all $x(k) \neq 0$.

The difference between corresponding trajectory of the system (4.35) is given by

$$\begin{aligned}
\mathbb{E}[\Delta V(x(k))] &= \mathbb{E}[V(x(k+1)) - V(x(k))] \\
&= \mathbb{E}[x^T(k+1)Px(k+1) - x^T(k)Px(k)] \\
&= \mathbb{E}[x^T(k)(A+BK)^T P(A+BK)x(k)] \\
&\quad\; + \mathbb{E}[x^T(k)A_p^T PA_p x(k)] - \mathbb{E}[x^T(k)Px(k)] \\
&= \mathbb{E}[x^T(k)[(A+BK)^T P(A+BK) + A_p^T PA_p - P]x(k),
\end{aligned} \tag{4.42}$$

which is stable if $\mathbb{E}[\Delta V(x(k))] < 0$. Then, we can form the discrete-time Lyapunov stabilizable equation as follows:

$$(A+BK)^T P(A+BK) + A_p^T PA_p - P = -(Q+2SK+K^T RK) \tag{4.43}$$

Substituting the value of $K = -\hat{R}^{-1}\left(B^T PA_S + R^{-1}S\right)$ to equation (4.43) and simplifying, we can get the following SARE:

$$P = A_s^T PA_s + A_p^T PA_p - A_s^T PB\hat{R}^{-1}B^T PA_s + Q_s, \tag{4.44}$$

where

$$\begin{aligned}
A_s &:= A - BR^{-1}S^T, \\
\hat{R} &:= R + B^T PB, \\
Q_s &:= Q - SR^{-1}S^T,
\end{aligned}$$

Hence, Lemma 4.3 is proved. □

4.3 Problem formulation

Consider a linear discrete-time stochastic system with state-dependent noise defined by

$$\begin{cases} x(k+1) = A(k)x(k) + \sum_{j=1}^{N} \left[B_{0j}(k)u_{0j}(k) + B_j(k)u_j(k)\right] \\ \qquad\qquad +D(k)v(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} C(k)x(k) \\ G_0(k)\boldsymbol{u}_0(k) \\ G_1(k)u_1(k) \\ \vdots \\ G_N(k)u_N(k) \end{bmatrix}, \; G_i^T(k)G_i(k) = I_{m_i},\; k \in \mathbf{N}_{T_f},\; i = 0,\; 1,\; \dots ,N, \end{cases} \tag{4.45}$$

where $x(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^n)$ represents the state vector, $\boldsymbol{u}_0(k) = \mathbf{col}\left[\ u_{01}(k) \quad \cdots \quad u_{0N}(k)\ \right] \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_0})$, $m_0 = \sum_{i=0}^{N} m_{0i}$, $i = 1,\ \dots ,N$ represents the leader's control input corresponding to i-th follower and $u_i(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_i})$, $i = 1,\ \dots ,N$ represents the i-th follower's control input, $v(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})$ represents the disturbance input and $z(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})$ represents the controlled output.

In many real-world control problems, most physical systems and processes include unmodeled uncertainties in the deterministic exogenous input $v(k)$, such as external disturbances. These introduce serious difficulties in the control and design of systems, in contrast with the stochastic perturbations due to the Wiener process. The H_∞ control method is a well-known approach to reducing the influence of these inputs and plays an important role in reducing the effect of such deterministic disturbances.

Throughout this Chapter, let $(\Omega,\ \mathscr{F},\ \mathbf{P})$ be a given filtered probability space, where $w(k)$ is a real-valued random variable defined in the filtered probability space, second-order process with $\mathbb{E}[w(k)] = 0$ and $\mathbb{E}[w(s)w(k)] = \delta_{sk}$ [Zhang et al. (2007), Zhang et al. (2008)]. For simplicity, we choose the closed-loop state feedback information structure for the leader and followers control. Moreover, it provides the advantage of complete control allowing us to access the control tools directly. In practical, we can see the problem of packet switch operating problem referred in [Saksena and Cruz (1985)] is solved by a state feedback control.

On the other hand, the linear quadratic cost functionals of the leader and followers, are given by

$$\begin{aligned} &J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots ,u_N,\ v) \\ &:= \sum_{k=0}^{T_f} \mathbb{E}\left[x^T(k)Q_0(k)x(k) + \sum_{j=1}^{N} \left\{u_{0j}^T(k)R_{00j}(k)u_{0j}(k) + u_j^T(k)R_{0j}(k)u_j(k)\right\}\right], \end{aligned} \tag{4.46a}$$

$$
\begin{aligned}
&J_i(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots,u_N,\ v) \\
&:= \sum_{k=0}^{T_f} \mathbb{E}\left[x^T(k)Q_i(k)x(k) + u_{0i}^T(k)R_{0ii}(k)u_{0i}(k) + u_i^T(k)R_{ii}(k)u_i(k)\right],
\end{aligned} \tag{4.46b}
$$

where

$$
\begin{aligned}
&Q_0(k) := C^T(k)C(k),\ R_{00i}(k) = R_{00i}^T(k) > 0,\ R_{0i}(k) = R_{0i}^T(k) \geq 0, \\
&Q_i(k) = Q_i^T(k) \geq 0,\ R_{0ii}(k) = R_{0ii}^T(k) \geq 0,\ R_{ii}(k) = R_{ii}^T(k) > 0,
\end{aligned}
$$

for $i = 1,\ \dots,N$ are matrices for any time step k.

Definition 4.3. *For one leader and N follower's team problem, suppose that* $J_0(\boldsymbol{u}_0,\ u_1,\ \dots,u_N)$ *is the leader's cost functional, where* $\boldsymbol{u}_0$ *is the leader's control and* $u_i,\ i = 1,\ 2,\ \dots,N$ *is the i-th follower's control. A strategy-set* $(\boldsymbol{u}_0^*,\ u_1^*,\ \dots,u_N^*)$ *is called the team-optimal solution of the game if*

$$
J_0(\boldsymbol{u}_0^*,\ u_1^*,\ \dots,u_N^*) \leq J_0(\boldsymbol{u}_0,\ u_1,\ \dots,u_N), \tag{4.47}
$$

for any $\boldsymbol{u}_0$ *and* $u_i,\ i = 1,\ 2,\ \dots,N.$

It should be noted that if J_0 is LQ form and strict convex, then a unique optimal solution exists [Başar and Olsder (1999)].

According to [Basar and Selbuz (1979)], [Mizukami and Wu (1988)], the framework of the incentive Stackelberg games can be described as follows:

(a) The player P_0 announces the following feedback pattern strategy in advance to the players P_i:

$$
u_{0i}(k) = u_{0i}(k, x(k), u_i(k)) = \eta_{0i}(k)x(k) + \eta_{ii}(k)u_i(k), \tag{4.48}
$$

where $\eta_{0i}(k) \in \mathbb{R}^{m_{0i} \times n}$ and $\eta_{ii}(k) \in \mathbb{R}^{m_{0i} \times m_i}$, $i = 1,\ \dots,N$ are discrete strategy parameter matrices.

(b) Each player P_i decides his/her own optimal strategy $u_i^*(k)$, $i = 1,\ \dots,N$ under the Nash equilibrium solution concept, considering the announced strategy of the player P_0.

(c) The player P_0 finalizes the incentive Stackelberg strategy

$$
u_{0i}^*(k) = u_{0i}^*(k, x(k), u_i^*(k)) = \eta_{0i}(k)x(k) + \eta_{ii}(k)u_i^*(k), \tag{4.49}
$$

for each player P_i, $i = 1,\ \dots,N$ so that the team-optimal solution can be achieved.

The presence of the external disturbance $v(k)$ affects the controlled output $z(k)$ through the state vector $x(k)$ which is measured by the perturbed operator $L_{T_f} : l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}) \rightarrow l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})$ [Zhang et al. (2007)] and its H_∞-norm is defined by

$$\|L_{T_f}\|_{H_\infty} := \sup_{\substack{v \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}), \\ v \neq 0,\ x_0 = 0}} \frac{\|z\|_{l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})}}{\|v\|_{l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})}}, \tag{4.50}$$

where

$$\|z\|_{l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_z})} := \left(x^T(k)Q_0(k)x(k) + \sum_{j=0}^{N} u_j^T(k)u_j(k) \right)^{\frac{1}{2}},$$

$$\|v\|_{l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v})} := \left(\sum_{k=0}^{T_f} \mathbb{E}\left[v^T(k)v(k)\right] \right)^{\frac{1}{2}}.$$

An important fact is that the effect of this disturbance cannot be avoided but weakened to some extent (disturbance attenuation level) $\gamma > \|L_{T_f}\|_{H_\infty}$. In other words, it is designed as the team controller $(\boldsymbol{u}_0,\ u_1,\ \dots,u_N)$ and the disturbance v are playing a zero-sum game, in which the cost is

$$J_v(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots,u_N,\ v) = \sum_{k=0}^{T_f} \mathbb{E}[\gamma^2\|v(k)\|^2 - \|z(k)\|^2], \quad \forall v(k) \neq 0. \tag{4.51}$$

In order to attenuate the efficiency of the disturbance under the H_∞-norm, the problem of H_∞-constraint is inevitable.

The finite-horizon H_∞-constrained incentive Stackelberg game with multiple non-cooperative followers can be formulated as follows. For any disturbance attenuation level $\gamma > 0,\ 0 < T_f < \infty$, we need to find an incentive strategy of P_0 by (4.49) and a state feedback strategy

$$u_i^*(k) := K_i(k)x(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{m_i}),$$

of $P_i,\ i = 1,\ \dots,N$ considering the worst-case disturbance

$$v^*(k) := F_\gamma(k)x(k) \in l_w^2(\mathbf{N}_{T_f},\ \mathbb{R}^{n_v}),$$

such that

(i) The trajectory of the closed-loop system (4.45) satisfies the following team-optimal condition (4.52a) along with H_∞ constraint condition (4.52b),

$$J_0(x_0,\boldsymbol{u}_0^*,u_1^*,\ \dots,u_N^*,v^*) \leq J_0(x_0,\boldsymbol{u}_0,u_1,\dots,u_N,v^*), \tag{4.52a}$$

$$0 \leq J_v(x_0,\boldsymbol{u}_0^*,u_1^*,\ \dots,u_N^*,v^*) \leq J_v(x_0,\boldsymbol{u}_0^*,u_1^*,\ \dots,u_N^*,v). \tag{4.52b}$$

(ii) A set of decision $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i=1,\ \ldots\ ,N$ satisfying the following Nash equilibrium inequality:

$$J_i(x_0,\boldsymbol{u}_0^*,u_1^*,\ \ldots\ ,u_N^*,v^*) \le J_i(x_0,\gamma_{-i}^*(u_{0i}),\gamma_{-i}^*(u_i),v^*). \tag{4.53}$$

Then, the strategy-set $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i=1,\ \ldots\ ,N$ constitutes both a team-optimal incentive Stackelberg strategy with the H_∞ constraint of the leader and Nash equilibrium strategies of the followers for a two-level hierarchical game [Başar and Olsder (1999)].

4.4 Main result

Suppose that for each strategy pair $(u_{0i},\ u_i) \in \Gamma_0 \times \Gamma_i$, the linear discrete-time stochastic system (4.45) has a unique solution for all x_0, and the value of J_i are well-defined, where Γ_i is defined as the space of admissible strategy of player $\mathrm{P}_i,\ i=0,\ 1,\ \ldots\ ,N$. First, to find the team-optimal solution triplet $(u_{0i}^*,\ u_i^*,\ v^*)$ with the H_∞ constraint, we centralize the control inputs of the system (4.45) as follows,

$$\begin{cases} x(k+1) = A(k)x(k)+B_c(k)u_c(k)+D(k)v(k)+A_p(k)x(k)w(k), \quad x(0)=x_0, \\ z(k) = \begin{bmatrix} C(k)x(k) \\ G_c(k)u_c(k) \end{bmatrix}, \quad G_c^T(k)G_c(k) = I_{\sum_{i=1}^N (m_{0i}+m_i)}, \quad k \in \mathbf{N}_{T_f}, \end{cases} \tag{4.54}$$

where

$$\begin{aligned} B_c(k) &:= \begin{bmatrix} \boldsymbol{B}_0(k) & B_1(k) & \cdots & B_N(k) \end{bmatrix}, \\ u_c(k) &:= \mathbf{col}\begin{bmatrix} \boldsymbol{u}_0(k) & u_1(k) & \cdots & u_N(k) \end{bmatrix}, \\ \boldsymbol{B}_0(k) &:= \begin{bmatrix} B_{01}(k) & \cdots & B_{0N}(k) \end{bmatrix}, \\ G_c(k) &:= \mathbf{block\ diag}\left(G_0(k)\ \ G_1(k)\ \cdots\ G_N(k)\right). \end{aligned}$$

Moreover, the leader's cost functional (4.46a) can be rewritten as,

$$J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \ldots\ ,u_N,\ v) := \sum_{k=0}^{T_f} \mathbb{E}\left[x^T(k)Q_0(k)x(k)+u_c^T(k)R_c(k)u_c(k)\right], \tag{4.55}$$

where

$$\begin{aligned} Q_0(k) &:= C^T(k)C(k), \\ R_c &:= \mathbf{block\ diag}\left(\boldsymbol{R}_{00}(k) \quad R_{01}(k) \quad \cdots \quad R_{0N}(k) \right), \\ \boldsymbol{R}_{00} &:= \mathbf{block\ diag}\left(R_{001}(k) \quad \cdots \quad R_{00N}(k) \right). \end{aligned}$$

Now, to apply Lemma 4.1, we assume that the following four cross-coupled SMVDEs have solutions $(P(k),\ W(k);K_c(k),\ F_\gamma(k))$ with $P(k) \geq 0$ and $W(k) \leq 0$:

$$
\begin{aligned}
P(k) &= A_v^T(k)P(k+1)A_v(k) + A_p^T(k)P(k+1)A_p - A_v^T(k)P(k+1)S_c(k)P(k+1)A_v(k) \\
&\quad + Q_0(k), \qquad P(T_f+1) = 0, && \text{(4.56a)}\\
K_c(k) &= -\hat{R}_c^{-1}(k)B^T(k)P(k+1)A_v(k), && \text{(4.56b)}\\
W(k) &= A_u^T(k)W(k+1)A_u(k) + A_p^T(k)W(k+1)A_p(k) - A_u^T(k)W(k+1)U(k)W(k+1)A_u(k) \\
&\quad - L_K(k), \qquad W(T_f+1) = 0, && \text{(4.56c)}\\
F_\gamma(k) &= -T_\gamma^{-1}(k)D^T(k)W(k+1)A_u(k), && \text{(4.56d)}
\end{aligned}
$$

where

$$
\begin{aligned}
&A_v(k) := A(k) + D(k)F_\gamma(k),\\
&\hat{R}_c(k) := I_{n_u} + B^T(k)P(k+1)B(k) > 0,\\
&K_c(k) := \begin{bmatrix} \boldsymbol{K}_0(k) \\ \boldsymbol{K}_1(k) \end{bmatrix},\\
&\boldsymbol{K}_0(k) := \begin{bmatrix} K_{01}^T(k) & \cdots & K_{0N}^T(k) \end{bmatrix}^T,\\
&\boldsymbol{K}_1(k) := \begin{bmatrix} K_1^T(k) & \cdots & K_N^T(k) \end{bmatrix}^T,\\
&A_u(k) := A(k) + B_c(k)K_c(k),\\
&L_K(k) := Q_0(k) + K_c(k)^T K_c(k),\\
&S_c(k) := B_c(k)\hat{R}_c^{-1}B_c^T(k),\\
&T_\gamma(k) := \gamma^2 I_{n_v} + D^T(k)W(k+1)D(k) > 0,\\
&U(k) := D(k)T_\gamma^{-1}(k)D^T(k).
\end{aligned}
$$

Then, we find the state feedback strategy pair

$$
(u_c^*(k),\ v^*(k)) := (K_c(k)x(k),\ F_\gamma(k)x(k)).
$$

This strategy pair is the team-optimal solution with the H_∞ constraint. More explicitly,

$$
\begin{aligned}
u_{0i}^*(k) &= K_{0i}(k)x(k)\\
&= -\left[R_{00i}(k) + B_{0i}^T(k)P(k+1)B_{0i}(k)\right]^{-1} B_{0i}^T(k)P(k+1)A_v(k)x(k), && \text{(4.57a)}\\
u_i^*(k) &= K_i(k)x(k)\\
&= -\left[R_{0i}(k) + B_i^T(k)P(k+1)B_i(k)\right]^{-1} B_i^T(k)P(k+1)A_v(k)x(k), && \text{(4.57b)}\\
v^*(k) &= F_\gamma(k)x(k)\\
&= -[\gamma^2 I_{n_v} + D^T(k)W(k+1)D(k)]^{-1}D^T(k)W(k+1)A_u(k)x(k). && \text{(4.57c)}
\end{aligned}
$$

It is assumed that the leader chooses the following incentive Stackelberg strategy corresponding to the i-th follower [Mizukami and Wu (1988)]:

$$u_{0i}^*(k) := \eta_{0i}(k)x(k) + \eta_{ii}(k)u_i^*(k), \qquad i = 1,\ 2,\ \dots,N. \tag{4.58}$$

Using (4.57a), (4.57b) and (4.58) $\eta_{0i}(k)$ and $\eta_{ii}(k)$ have the following relation:

$$\eta_{0i}(k) := K_{0i}(k) - \eta_{ii}(k)K_i(k), \qquad i = 1,\ 2,\ \dots,N. \tag{4.59}$$

As the second step, the non-cooperative followers' Nash strategy-set is derived where the leader's incentive Stackelberg strategy is considered. For this purpose, we can rewrite the system for i-th follower using the strategy triplet $(\gamma_{-i}^*(u_{0i}),\ \gamma_{-i}^*(u_i),\ v^*)$ as follows,

$$\begin{aligned} x(k+1) =& A(k)x(k) + \sum_{j=1,\ j\neq i}^{N} B_{0j}(k)u_{0j}^*(k) + B_{0i}(k)[\eta_{0i}(k)x(k) + \eta_{ii}(k)u_i(k)] \\ &+ \sum_{j=1,\ j\neq i}^{N} B_j(k)u_j^*(k) + B_i(k)u_i(k) + D(k)v^*(k) + A_p(k)x(k)w(k),\ x(0) = x_0, \end{aligned} \tag{4.60}$$

with cost functional

$$\begin{aligned} & J_i(x_0,\ \gamma_{-i}^*(u_{0i}),\ \gamma_{-i}^*(u_i),\ v^*) \\ & := \sum_{k=0}^{T_f} \mathbb{E}\Big[x^T(k)Q_i(k)x(k) + [\eta_{0i}(k)x(k) + \eta_{ii}(k)u_i(k)]^T R_{0ii}(k)[\eta_{0i}(k)x(k) \\ & \quad + \eta_{ii}(k)u_i(k)] + u_i^T(k)R_{ii}(k)u_i(k)\Big], \qquad i = 1,\ \dots,N. \end{aligned} \tag{4.61}$$

Using simplified notations the above system can be written as

$$x(k+1) = \tilde{A}_i(k)x(k) + \tilde{B}_i(k)u_i(k) + A_p(k)x(k)w(k), \quad x(0) = x_0, \tag{4.62}$$

with cost functional

$$\begin{aligned} & J_i(x_0,\ \gamma_{-i}^*(u_{0i}),\ \gamma_{-i}^*(u_i),\ v^*) \\ & := \sum_{k=0}^{T_f} \mathbb{E}\Big[x^T(k)\tilde{Q}_i(k)x(k) + 2x_i^T(k)\tilde{S}_i(k)u_i(k) + u_i^T(k)\tilde{R}_i(k)u_i(k)\Big], \quad i = 1,\ \dots,N, \end{aligned} \tag{4.63}$$

where

$$\tilde{A}_i(k) := A(k) + D(k)F_\gamma(k) + B_{0i}(k)\eta_{0i}(k) + \sum_{j=1,\ j\neq i}^{N} [B_{0j}(k)K_{0j}(k) + B_j(k)K_j(k)],$$

$$\tilde{B}_i(k) := B_i(k) + B_{0i}(k)\eta_{ii}(k),$$
$$\tilde{S}_i(k) := \eta_{0i}^T(k)R_{0ii}(k)\eta_{ii}(k),$$
$$\tilde{Q}_i(k) := Q_i(k) + \eta_{0i}^T(k)R_{0ii}(k)\eta_{0i}(k),$$
$$\tilde{R}_i(k) := R_{ii}(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{ii}(k).$$

Now, by applying Lemma 4.2, the following SBDRE has solution(s) $P_i(k)$ corresponding to each i-th follower:

$$\begin{aligned} P_i(k) =& \hat{A}_i^T(k)P_i(k+1)\hat{A}_i(k) + A_p^T(k)P_i(k+1)A_p(k) - \hat{A}_i^T(k)P_i(k+1)\tilde{B}_i(k) \\ & \times \hat{R}_i^{-1}(k)\tilde{B}_i^T(k)P_i(k+1)\hat{A}_i(k) + \hat{Q}_i(k), \quad P_i(T_f+1) = 0, \end{aligned} \tag{4.64}$$

where

$$\hat{A}_i(k) := \tilde{A}_i(k) - \tilde{B}_i(k)\tilde{R}_i^{-1}(k)\tilde{S}_i^T(k),$$
$$\hat{R}_i(k) := \tilde{R}_i(k) + \tilde{B}_i^T(k)P_i(k+1)\tilde{B}_i(k),$$
$$\hat{Q}_i(k) := \tilde{Q}_i(k) - \tilde{S}_i(k)\tilde{R}_i^{-1}(k)\tilde{S}_i^T(k),$$

and each i-th follower's optimal state feedback Nash equilibrium strategy is determined by

$$u_i^*(k) = \tilde{K}_i(k)x(k) = -\hat{R}_i^{-1}(k)\left[\tilde{S}_i^T(k) + \tilde{B}_i^T(k)P_i(k+1)\hat{A}_i(k)\right]x(k), \tag{4.65}$$

$i = 1, \ldots, N$. Owing to the equivalence of (4.57b) and (4.65), that is $K_i(k) = \tilde{K}_i(k)$ we have the relation,

$$K_i(k) = -\hat{R}_i^{-1}(k)\left[\tilde{S}_i^T(k) + \tilde{B}_i^T(k)P_i(k+1)\tilde{A}_i\right],$$

or,

$$[\tilde{R}_i(k) + \tilde{B}_i^T(k)P_i(k+1)\tilde{B}_i(k)]K_i(k) = -\tilde{S}_i^T(k) - \tilde{B}_i^T(k)P_i(k+1)\tilde{A}_i(k),$$

or,

$$\begin{aligned} &[R_{ii}(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{ii}(k) + [B_i(k) + B_{0i}(k)\eta_{ii}(k)]^T P_i(k+1)\tilde{B}_i(k)]K_i(k) \\ &\quad = -[\eta_{0i}^T(k)R_{0ii}(k)\eta_{ii}(k)]^T - [B_i(k) + B_{0i}(k)\eta_{ii}(k)]^T P_i(k+1)\tilde{A}_i(k), \end{aligned}$$

or,

$$\begin{aligned} &R_{ii}(k)K_i(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{ii}(k)K_i(k) \\ &\quad + B_i^T(k)P_i(k+1)\tilde{B}_i(k)K_i(k) + \eta_{ii}^T(k)B_{0i}^T(k)P_i(k+1)\tilde{B}_i(k)K_i(k) \\ &\quad = -\eta_{ii}^T(k)R_{0ii}(k)K_{0i}(k) + \eta_{ii}^T(k)R_{0ii}(k)\eta_{ii}(k)K_i(k) - B_i^T(k)P_i(k+1)\tilde{A}_i(k) \end{aligned}$$

$$- \eta_{ii}^T(k) B_{0i}^T(k) P_i(k+1) \tilde{A}_i(k),$$
$$[\text{using relation (4.59) } \eta_{0i}(k) = K_{0i}(k) - \eta_{ii}(k) K_i(k)],$$

or,

$$\eta_{ii}^T(k) = -[R_{ii}(k)K_i(k) + B_i^T(k)P_i(k+1)\tilde{B}_i(k)K_i(k) + B_i(k)^T(k)P_i(k+1)\tilde{A}_i(k)]$$
$$\times [B_{0i}^T(k)P_i(k+1)\tilde{B}_i(k)K_i(k) + B_{0i}^T P_i(k+1)\tilde{A}_i(k) + R_{0ii}(k)K_{0i}(k)]^{-1}$$

or,

$$\eta_{ii}^T(k) = -\left(R_{ii}(k)K_i(k) + B_i^T(k)P_i(k+1)\tilde{\mathbb{A}}_i(k)\right)\left(R_{0ii}(k)K_{0i}(k) + B_{0i}^T(k)P_i(k+1)\tilde{\mathbb{A}}_i(k)\right)^{-1}, \tag{4.66}$$

where

$$\tilde{\mathbb{A}}_i(k) := \tilde{A}_i(k) + \tilde{B}_i(k)K_i(k), \qquad i = 1, \ldots, N.$$

Remark 4.1. *It should be noted that the incentive parameter* $\eta_{ii}(k)$ *can be uniquely determined if and only if* $\left(R_{0ii}(k)K_{0i}(k) + B_{0i}^T(k)P_i(k+1)\tilde{\mathbb{A}}_i(k)\right)$ *is non-singular.*

Theorem 4.1. *Suppose that four cross-coupled SMVDEs* (4.56)*, SBDRE* (4.64)*, and equation* (4.66) *have solutions. Then, strategy-set* (4.49) *is associated with* (4.65) *from the two-level incentive Stackelberg strategy-set with* H_∞ *constraint as formulated in Section 4.4, where* η_{0i} *and* η_{ii} *are determined through* (4.59) *and* (4.66)*, respectively.*

Remark 4.2. *It should be noted that if we substitute* $\eta_{0i}(k)$ *from relation* (4.59) *into SBDRE* (4.64)*, then SBDRE* (4.64) *will have two unknowns,* $P_i(k)$ *and* $\eta_{ii}(k)$*. Further, by solving the two equations,* (4.64) *and* (4.66)*, it will be possible to obtain a solution for* $P_i(k)$ *and* $\eta_{ii}(k)$*.*

4.5 Infinite-horizon Case

The infinite-horizon H_∞-constrained incentive Stackelberg game is investigated in this section. Consider a time-invariant linear stochastic discrete-time system such as,

$$\begin{cases} x(k+1) = Ax(k) + \sum_{j=1}^{N} B_{0j}u_{0j}(k) + \sum_{j=1}^{N} B_j u_j(k) \\ \qquad + Dv(k) + A_p x(k) w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} Cx(k) \\ G_0 \boldsymbol{u}_0(k) \\ G_1 u_1(k) \\ \vdots \\ G_N u_N(k) \end{bmatrix}, \quad G_i^T G_i = I_{n_u}, \quad k \in \mathbf{N}. \end{cases} \tag{4.67}$$

Moreover, the cost functionals are defined as

$$
\begin{aligned}
&J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \ldots,u_N,\ v) \\
&\quad := \sum_{k=0}^{\infty} \mathbb{E}\left[x^T(k)Q_0x(k) + \sum_{j=1}^{N}\left\{u_{0j}^T(k)R_{00j}u_{0j}(k) + u_j^T(k)R_{0j}u_j(k)\right\}\right], \qquad (4.68a)\\
&J_i(x_0,\ \boldsymbol{u}_0,\ u_1,\ \ldots,u_N,\ v) \\
&\quad := \sum_{k=0}^{\infty} \mathbb{E}\left[x^T(k)Q_ix(k) + \sum_{j=1}^{N} u_{0j}^T(k)R_{0ij}u_{0j}(k) + u_i^T(k)R_{ii}u_i(k)\right],\ i = 1,\ \ldots,N, \qquad (4.68b)
\end{aligned}
$$

where

$$
\begin{aligned}
&Q_0 := C^TC,\ R_{00i} = R_{00i}^T > 0,\ R_{0i} = R_{0i}^T \geq 0,\\
&Q_i = Q_i^T \geq 0,\ R_{0ii} = R_{0ii}^T \geq 0,\ R_{ii} = R_{ii}^T > 0,
\end{aligned}
$$

for all $i = 0,\ 1,\ \ldots,N$. The infinite-horizon H_∞-constrained incentive Stackelberg game with multiple non-cooperative followers can be formulated as follows:

For any disturbance attenuation level $\gamma > 0$, we need to find an incentive strategy of P_0 by

$$
u_{0i}^*(k) = u_{0i}^*(k, x(k), u_i^*(k)) = \eta_{0i}x(k) + \eta_{ii}u_i^*(k), \qquad (4.69)
$$

where parameters η_{0i} and η_{ii} are to be determined and a state feedback strategy

$$
u_i^*(k) := K_ix(k) \in l_w^2(\mathbf{N},\ \mathbb{R}^{m_i}),
$$

of $\mathrm{P}_i,\ i = 1,\ \ldots,N$ considering the worst-case disturbance

$$
v^*(k) = F_\gamma x(k) \in l_w^2(\mathbf{N},\ \mathbb{R}^{n_v}),
$$

such that

(i) the trajectory of the closed-loop system (4.67) satisfies the following team-optimal condition (4.70a) along with H_∞ constraint conditions (4.70b)

$$
\begin{aligned}
&J_0(x_0, \boldsymbol{u}_0^*, u_1^*,\ \ldots,u_N^*, v^*) \leq J_0(x_0,\ \boldsymbol{u}_0, u_1,\ \ldots,u_N, v^*), \qquad (4.70a)\\
&0 \leq J_v(x_0, \boldsymbol{u}_0^*, u_1^*,\ \ldots,u_N^*, v^*) \leq J_v(x_0, \boldsymbol{u}_0^*, u_1^*,\ \ldots,u_N^*, v), \qquad (4.70b)
\end{aligned}
$$

where

$$
\begin{aligned}
&J_v(x_0,\ \boldsymbol{u}_0,\ u_1,\ \ldots,u_N,\ v) := \sum_{k=0}^{T_f} \mathbb{E}[\gamma^2\|v(k)\|^2 - \|z(k)\|^2], \quad \forall v(k) \neq 0,\\
&\|z(k)\|^2 = x^T(k)Q_0x(k) + \sum_{j=0}^{N} u_j^T(k)u_j(k),
\end{aligned}
$$

(ii) a set of decision $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i=1,\ \dots\ ,N$ satisfying the following Nash equilibrium inequality:

$$J_i(x_0, \boldsymbol{u}_0^*, u_1^*,\ \dots\ , u_N^*, v^*) \le J_i(x_0, \gamma_{-i}^*(u_{0i}), \gamma_{-i}^*(u_i), v^*). \tag{4.71}$$

Then, the strategy-set $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i=1,\ \dots\ ,N$ constitutes both team-optimal incentive Stackelberg strategy with H_∞ constraint of the leader and Nash equilibrium strategies of the followers for a two-level hierarchical game [Başar and Olsder (1999)].

Note that, if the inequality (4.70b) holds, it ensures the following condition for H_∞ constraint:

$$\|L\|_{H_\infty} = \sup_{\substack{v \in l_w^2(\mathbf{N},\ \mathbb{R}^{n_v}),\\ v \neq 0,\ x_0 = 0}} \frac{\|z\|_{l_w^2(\mathbf{N},\ \mathbb{R}^{n_z})}}{\|v\|_{l_w^2(\mathbf{N},\ \mathbb{R}^{n_v})}} < \gamma. \tag{4.72}$$

First, to find the team-optimal solution triplet $(u_{0i}^*,\ u_i^*,\ v^*)$ with H_∞ constraint, rearrange the system (4.67) as follows,

$$\begin{cases} x(k+1) = Ax(k) + B_c u_c(k) + Dv(k) + A_p x(k) w(k), \quad x(0) = x_0, \\ z(k) = \begin{bmatrix} Cx(k) \\ G_c u_c(k) \end{bmatrix},\ G_c^T G_c = I_{\sum_{i=1}^N (m_{0i}+m_i)},\ k \in \mathbf{N}, \end{cases} \tag{4.73}$$

where $x(0) = x_0,\ k \in \mathbf{N}$,

$$\begin{aligned} B_c &:= \begin{bmatrix} \boldsymbol{B}_0 & B_1 & \cdots & B_N \end{bmatrix}, \\ \boldsymbol{B}_0 &:= \begin{bmatrix} B_{01} & \cdots & B_{0N} \end{bmatrix}, \\ u_c(k) &:= \mathbf{col} \begin{bmatrix} \boldsymbol{u}_0(k) & u_1(k) & \cdots & u_N(k) \end{bmatrix}, \\ G_c &:= \mathbf{block\ diag} \begin{pmatrix} G_0 & G_1 & \cdots & G_N \end{pmatrix}. \end{aligned}$$

Moreover, the leader's cost functional (4.68a) can be rewritten as

$$J_0(x_0,\ \boldsymbol{u}_0,\ u_1,\ \dots\ , u_N,\ v) := \sum_{k=0}^{\infty} \mathbb{E}\left[x^T(k) Q_0 x(k) + u_c^T(k) R_c u_c(k)\right], \tag{4.74}$$

where

$$\begin{aligned} R_c &:= \mathbf{block\ diag} \begin{pmatrix} \boldsymbol{R}_{00} & R_{01} & \cdots & R_{0N} \end{pmatrix}, \\ \boldsymbol{R}_{00} &:= \mathbf{block\ diag} \begin{pmatrix} R_{001} & \cdots & R_{00N} \end{pmatrix}. \end{aligned}$$

Now, according to [Zhang et al. (2008)], for the system (4.73), suppose the following four cross-coupled SMVAEs have solutions $(P,\ W;\ K_c,\ F_\gamma)$ with $P > 0$ and $W < 0$:

$$P = A_v^T P A_v + A_p^T P A_p - A_v^T P S_c P A_v + Q_0, \tag{4.75a}$$

$$K_c = -\hat{R}_c^{-1} B^T P A_v, \tag{4.75b}$$
$$W = A_u^T W A_u + A_p^T W A_p - A_u^T W U W A_u - L_K, \tag{4.75c}$$
$$F_\gamma = -T_\gamma^{-1} D^T W A_u, \tag{4.75d}$$

where

$$\begin{aligned}
A_v &:= A + DF_\gamma, \\
\hat{R}_c &:= I_{n_u} + B^T P B > 0, \\
A_u &:= A + B_c K_c, \\
L_K &:= Q_0 + K_c^T K_c, \\
K_c &:= \begin{bmatrix} \boldsymbol{K}_0 \\ \boldsymbol{K}_1 \end{bmatrix}, \\
\boldsymbol{K}_0 &:= \begin{bmatrix} K_{01}^T & \cdots & K_{0N}^T \end{bmatrix}^T, \\
\boldsymbol{K}_1 &:= \begin{bmatrix} K_1^T & \cdots & K_N^T \end{bmatrix}^T, \\
S_c &:= B_c \hat{R}_c^{-1} B_c^T, \\
U &:= D T_\gamma^{-1} D^T, \\
T_\gamma &:= \gamma^2 I_{n_v} + D^T W D > 0.
\end{aligned}$$

If $(A,\ A_p/C)$ and $(A + DF_\gamma,\ A_p/C)$ are exactly observable, then the state feedback strategy pair

$$(u_c^*(k),\ v^*(k)) := (K_c x(k),\ F_\gamma x(k)),$$

is the team-optimal solution for the system (4.73) under H_∞ constraint. More explicitly,

$$u_{0i}^*(k) = K_{0i} x(k) = -\left[R_{00i} + B_{0i}^T P B_{0i}\right]^{-1} B_{0i}^T P A_v x(k), \tag{4.76a}$$
$$u_i^*(k) = K_i x(k) = -\left[R_{0i} + B_i^T P B_i\right]^{-1} B_i^T P A_v x(k), \tag{4.76b}$$
$$v^*(k) = F_\gamma x(k) = -[\gamma^2 I_{n_v} + D^T W D]^{-1} D^T W A_u x(k). \tag{4.76c}$$

Secondly, to derive followers' Nash strategy-set, we can rewrite the system for i-th follower as

$$x(k+1) = \tilde{A}_i x(k) + \tilde{B}_i u_i(k) + A_p x(k) w(k),\ x(0) = x_0, \tag{4.77}$$

with cost functional

$$J_i(x_0,\ \gamma_{-i}^*(u_{0i}),\ \gamma_{-i}^*(u_i),\ v^*) := \sum_{k=0}^{\infty} \mathbb{E}\left[x^T(k)\tilde{Q}_i x(k) + 2x_i^T \tilde{S}_i u_i(k) + u_i^T(k)\tilde{R}_i u_i(k)\right], \tag{4.78}$$

where

$$\tilde{A}_i := A + DF_\gamma + B_{0i}\eta_{0i} + \sum_{j=1,\ j\neq i}^{N} [B_{0j}K_{0j} + B_j K_j],$$

$$\tilde{B}_i := B_i + B_{0i}\eta_{ii},\ \tilde{Q}_i := Q_i + \eta_{0i}^T R_{0ii}\eta_{0i},$$
$$\tilde{S}_i := \eta_{0i}^T R_{0ii}\eta_{ii},$$
$$\tilde{R}_i := R_{ii} + \eta_{ii}^T R_{0ii}\eta_{ii}, \quad i = 1,\ \dots ,N.$$

Now, applying Lemma 4.3, there exists a symmetric constant matrix $P_i > 0$ corresponding to each i-th follower that solves the following SARE:

$$P_i = \hat{A}_i^T P_i\hat{A}_i + A_p^T P_i A_p - \hat{A}_i^T P_i\tilde{B}_i\hat{R}_i^{-1}\tilde{B}_i^T P_i\hat{A}_i + \hat{Q}_i, \tag{4.79}$$

where

$$\hat{A}_i := \tilde{A}_i - \tilde{B}_i\tilde{R}_i^{-1}\tilde{S}_i^T,\ \hat{R}_i := \tilde{R}_i + \tilde{B}_i^T P_i\tilde{B}_i,\ \hat{Q}_i := \tilde{Q}_i - \tilde{S}_i\tilde{R}_i^{-1}\tilde{S}_i^T.$$

Then, each i-th follower's optimal state feedback Nash equilibrium strategy is determined by

$$u_i^*(k) = \tilde{K}_i x(k) = -\hat{R}_i^{-1}\left[\tilde{S}_i^T + \tilde{B}_i^T P_i\hat{A}_i\right] x(k),\ i = 1,\ \dots ,N. \tag{4.80}$$

Due to the equivalence of (4.76b) and (4.80), that is $K_i = \tilde{K}_i$ we have the relation

$$K_i = -\hat{R}_i^{-1}\left[\tilde{S}_i^T + \tilde{B}_i^T P_i\tilde{A}_i\right],$$

or,

$$[\tilde{R}_i + \tilde{B}_i^T P_i\tilde{B}_i]K_i = -\tilde{S}_i^T - \tilde{B}_i^T P_i\tilde{A}_i,$$

or,

$$[R_{ii} + \eta_{ii}^T R_{0ii}\eta_{ii} + [B_i + B_{0i}\eta_{ii}]^T P_i\tilde{B}_i]K_i = -[\eta_{0i}^T R_{0ii}\eta_{ii}]^T - [B_i + B_{0i}\eta_{ii}]^T P_i\tilde{A}_i,$$

or,

$$\begin{aligned} &R_{ii}K_i + \eta_{ii}^T R_{0ii}\eta_{ii}K_i + B_i^T P_i\tilde{B}_i K_i + \eta_{ii}^T B_{0i}^T P_i\tilde{B}_i K_i \\ &\quad = -\eta_{ii}^T R_{0ii}K_{0i} + \eta_{ii}^T R_{0ii}\eta_{ii}K_i - B_i^T P_i\tilde{A}_i - \eta_{ii}^T B_{0i}^T P_i\tilde{A}_i, \\ &\qquad [\text{using relation } \eta_{0i} = K_{0i} - \eta_{ii}K_i] \end{aligned}$$

or,

$$\eta_{ii}^T = -[R_{ii}K_i + B_i^T P_i\tilde{B}_i K_i + B_i^T P_i\tilde{A}_i][B_{0i}^T P_i\tilde{B}_i K_i + B_{0i}^T P_i\tilde{A}_i + R_{0ii}K_{0i}]^{-1},$$

or,

$$\eta_{ii}^T = -\left(R_{ii}K_i + B_i^T P_i\tilde{\mathbb{A}}_i\right)\left(R_{0ii}K_{0i} + B_{0i}^T P_i\tilde{\mathbb{A}}_i\right)^{-1}, \tag{4.81}$$

where

$$\tilde{\mathbb{A}}_i := \tilde{A}_i + \tilde{B}_i K_i, \qquad i = 1,\ \dots ,N.$$

Remark 4.3. *It should be noted that the incentive parameter* η_{ii} *can be uniquely determined if and only if* $\left(R_{0ii}K_{0i}+B_{0i}^TP_i\tilde{\mathbb{A}}_i\right)$ *is non-singular.*

Parameter η_{0i} can be determined by the following equation:

$$\eta_{0i}=K_{0i}-\eta_{ii}K_i, \qquad i=1,\ 2,\ \dots,N. \tag{4.82}$$

Theorem 4.2. *Suppose that a discrete-time stochastic system* (4.67) *is stabilizable and that four cross-coupled SMVAEs* (4.75) *have the solution set* $(P,\ W;\ K_c,\ F_\gamma)$ *such that* $P>0$, $W<0$ *and* $(A,\ A_p/C)$ *and* $(A+DF_\gamma,\ A_p/C)$ *are exactly observable. If SARE* (4.79) *and equation* (4.81) *have solutions* $P_i>0$ *and* η_{ii}*, respectively, then the strategy-sets* (4.69) *and* (4.80) *from the two-level incentive Stackelberg strategy-set with* H_∞ *constraint are formulated, as shown in Section 4.5.*

In order to solve four cross-coupled SMVAEs of (4.75) and SARE (4.79) along with (4.81) the following Lyapunov based computational algorithm is used:

$$\begin{cases} P^{(r+1)}=[A_v^{(r)}]^TP^{(r+1)}A_v^{(r)}+A_p^TP^{(r)}A_p-[A_v^{(r)}]^TP^{(r)}S_c^{(r)}P^{(r)}A_v^{(r)}+Q_0, \\ K_c^{(r)}=-[\hat{R}_c^{(r)}]^{-1}B_c^TP^{(r)}A_v^{(r)}, \\ W^{(r+1)}=[A_u^{(r)}]^TW^{(r+1)}A_u^{(r)}+A_p^TW^{(r)}A_p-[A_u^{(r)}]^TW^{(r)}U^{(r)}W^{(r)}A_u^{(r)}-L_K^{(r)}, \\ F_\gamma^{(r)}=-[T_\gamma^{(r)}]^{-1}D^TW^{(r)}A_u^{(r)}, \end{cases} \tag{4.83a}$$

$$\begin{cases} P_i^{(s+1)}=[\hat{A}_i^{(s)}]^TP_i^{(s+1)}\hat{A}_i^{(s)}+A_p^TP_i^{(s)}A_p-[\hat{A}_i^{(s)}]^TP_i^{(s)}\tilde{B}_i^{(s)}[\hat{R}_i^{(s)}]^{-1}[\tilde{B}_i^{(s)}]^TP_i^{(s)}\hat{A}_i^{(s)}+\hat{Q}_i^{(s)}, \\ [\eta_{ii}^{(s+1)}]^T=-\left(R_{ii}K_i^{(s)}+B_i^TP_i^{(s+1)}\tilde{\mathbb{A}}_i^{(s)}\right)\left(R_{0ii}K_{0i}+B_{0i}^TP_i^{(s+1)}\tilde{\mathbb{A}}_i^{(s)}\right)^{-1}, \end{cases} \tag{4.83b}$$

where $r=0,\ 1,\ \dots,\ s=0,\ 1,\ \dots,$

$$\begin{aligned} &P^{(0)}=P_i^{(0)}=I_n,\ W^{(0)}=-I_n,\ \eta_{ii}^{(0)}=\eta_{ii}^0, \\ &A_v^{(r)}:=A+DF_\gamma^{(r)}, \\ &A_u^{(r)}:=A+B_cK_c^{(r)}, \\ &S_c^{(r)}:=B_c[\hat{R}_c^{(r)}]^{-1}B_c^T, \\ &U^{(r)}:=D[T_\gamma^{(r)}]^{-1}D^T, \\ &T_\gamma^{(r)}:=\gamma^2I_{n_v}+D^TW^{(r)}D, \\ &\hat{R}_c^{(r)}:=R_c+B_c^TP^{(r)}B_c, \\ &L_K^{(r)}:=C^TC+[K_c^{(r)}]^TK_c^{(r)}, \\ &\tilde{A}_i^{(s)}:=A+DF_\gamma+B_{0i}\eta_{0i}^{(s)}+\sum_{j=1,\ j\neq i}^{N}[B_{0j}K_{0j}+B_jK_j], \\ &\tilde{B}_i^{(s)}:=B_i+B_{0i}\eta_{ii}^{(s)}, \end{aligned}$$

$$
\begin{aligned}
\tilde{S}_i &:= [\eta_{0i}^{(s)}]^T R_{0ii} \eta_{ii}^{(s)}, \\
\tilde{Q}_i^{(s)} &:= Q_i + [\eta_{0i}^{(s)}]^T R_{0ii} \eta_{0i}^{(s)}, \\
\tilde{R}_i^{(s)} &:= R_{ii} + [\eta_{ii}^{(s)}]^T R_{0ii} \eta_{ii}^{(s)}, \\
\tilde{\mathbb{A}}_i^{(s)} &:= \tilde{A}_i^{(s)} + \tilde{B}_i^{(s)} K_i^{(s)}, \\
\hat{A}_i^{(s)} &:= \tilde{A}_i^{(s)} - \tilde{B}_i^{(s)} [\tilde{R}_i^{(s)}]^{-1} [\tilde{S}_i^{(s)}]^T, \\
\hat{R}_i^{(s)} &:= \tilde{R}_i^{(s)} + [\tilde{B}_i^{(s)}]^T P_i \tilde{B}_i^{(s)}, \\
\hat{Q}_i^{(s)} &:= \tilde{Q}_i^{(s)} - \tilde{S}_i^{(s)} [\tilde{R}_i^{(s)}]^{-1} [\tilde{S}_i^{(s)}]^T \geq 0.
\end{aligned}
$$

It should be noted that the initial choice of η_{ii}^0 has to be chosen appropriately. In the next section, an academic and a practical numerical examples demonstrate that this algorithm operates well in practice.

4.6 Numerical examples

In this section, we investigate two simple numerical examples to demonstrate the existence of our proposed incentive Stackelberg strategy-set.

4.6.1 An academic example

First, we present an example for the infinite-horizon case with two non-cooperative players. Later, to show the convergence of our results, we present some graphs of the same problem in Fig. 4.2–4.5 considering a finite time interval. Let us consider the following system matrices:

$$
x(0) = \begin{bmatrix} 1 \\ 0.5 \\ -0.5 \end{bmatrix}, \; A = \begin{bmatrix} 0.52 & 1.12 & 0 \\ 0 & -0.24 & 0 \\ 0.23 & 0.85 & -0.16 \end{bmatrix}, \; A_p = 0.1A,
$$

$$
B_{01} = \begin{bmatrix} 0.138 & 0.20 & 1.15 \\ -0.55 & 0.84 & -1.11 \\ 5.23 & 0 & 0.11 \end{bmatrix}, \; B_{02} = \begin{bmatrix} 0.312 & 1.20 & 0.24 \\ -1.25 & 1.03 & 0.65 \\ 3.55 & 0 & 0.22 \end{bmatrix},
$$

$$
B_1 = \begin{bmatrix} 0.15 & -0.11 & 0.45 \\ 0.12 & 2.28 & 0.03 \\ 3.55 & 0 & 0.22 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0.23 & -0.45 & 0.22 \\ -0.52 & 1.02 & 0.02 \\ 0.28 & 2.11 & 1.96 \end{bmatrix},
$$

$$
D = \begin{bmatrix} 0.054 & -0.076 & 0.23 \\ -0.035 & -0.094 & 0.043 \\ 0.023 & 0.043 & 0.013 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix},
$$

$$
Q_0 = \mathbf{diag}(\; 1 \quad 1 \quad 2 \;), \quad Q_1 = \mathbf{diag}(\; 1 \quad 1.5 \quad 2.1 \;), \; Q_2 = \mathbf{diag}(\; 1.2 \quad 1.1 \quad 3.1 \;),
$$

$$
R_{001} = 1.9I_3, \quad R_{002} = 2.5I_3, \quad R_{01} = 2.7I_3, \quad R_{02} = 3.5I_3,
$$

$R_{011} = 4.8I_3, \quad R_{022} = 5I_3, \quad R_{11} = 0.3I_3, \quad R_{22} = 0.5I_3.$

We choose the disturbance attenuation level as $\gamma = 5$.

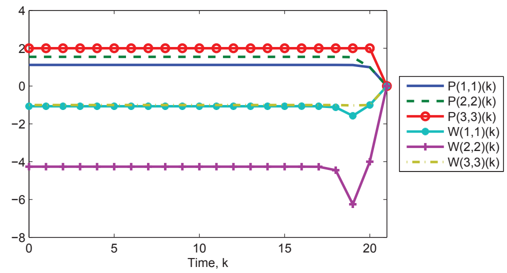


Fig. 4.2: Convergence graph of diagonal elements of $P(k)$, $W(k)$.

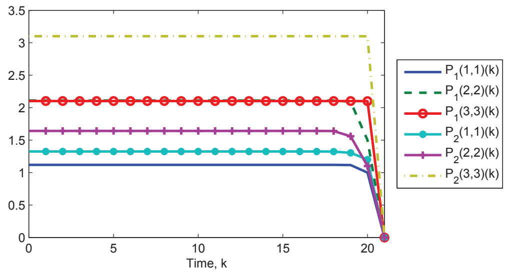


Fig. 4.3: Convergence graph of diagonal elements of $P_1(k)$, $P_2(k)$.

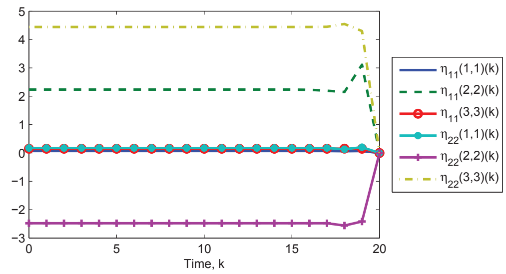


Fig. 4.4: Convergence graph of diagonal elements of $\eta_{11}(k)$, $\eta_{22}(k)$.

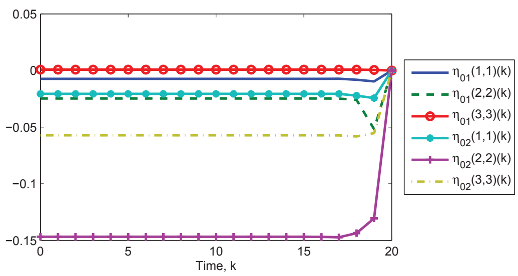


Fig. 4.5: Convergence graph of diagonal elements of $\eta_{01}(k)$, $\eta_{02}(k)$.

First, for the infinite-horizon case, four cross-coupled SMVAEs (4.75) are solved using algorithm (4.83a). These solutions attain an H_∞-constrained team-optimal solutions as presented below:

$$P = \begin{bmatrix} 1.1166 & 2.4780\text{e-}1 & -1.1222\text{e-}3 \\ 2.4780\text{e-}1 & 1.5448 & -5.6224\text{e-}3 \\ -1.1222\text{e-}3 & -5.6224\text{e-}3 & 2.0016 \end{bmatrix},$$

$$K_c = \begin{bmatrix} K_{01} \\ K_{02} \\ K_1 \\ K_2 \end{bmatrix} = \left[\begin{array}{ccc} -1.5655\text{e-}2 & -8.0571\text{e-}2 & 1.8506\text{e-}2 \\ -3.2647\text{e-}2 & -4.9939\text{e-}2 & 4.9357\text{e-}4 \\ -1.1706\text{e-}1 & -2.6894\text{e-}1 & -1.9856\text{e-}3 \\ \hline -1.8913\text{e-}2 & -7.8576\text{e-}2 & 8.7567\text{e-}3 \\ -1.1254\text{e-}1 & -2.1784\text{e-}1 & -4.3089\text{e-}4 \\ -2.6742\text{e-}2 & -4.6254\text{e-}2 & 8.2324\text{e-}4 \\ \hline -1.3440\text{e-}2 & -2.8877\text{e-}2 & 1.3626\text{e-}3 \\ -1.1534\text{e-}2 & 4.2314\text{e-}3 & 4.9096\text{e-}3 \\ -3.7172\text{e-}2 & -8.2997\text{e-}2 & 2.7749\text{e-}3 \\ \hline -1.7687\text{e-}2 & -3.1838\text{e-}2 & 6.5192\text{e-}4 \\ 1.9603\text{e-}2 & 4.4974\text{e-}2 & 4.9832\text{e-}3 \\ -1.4987\text{e-}2 & -3.8356\text{e-}2 & 3.7607\text{e-}3 \end{array}\right],$$

$$W = \begin{bmatrix} -1.0656 & -2.1303 & -9.9778\text{e-}1 \\ -2.1303 & -4.2664 & -1.9946 \\ -9.9778\text{e-}1 & -1.9946 & -1.0008 \end{bmatrix},$$

$$F_\gamma = \begin{bmatrix} 3.9128\text{e-}5 & 7.2135\text{e-}5 & -1.1007\text{e-}6 \\ -1.5614\text{e-}3 & -2.8465\text{e-}3 & 3.5938\text{e-}5 \\ 2.2915\text{e-}3 & 4.1800\text{e-}3 & -5.3389\text{e-}5 \end{bmatrix}.$$

Then, SARE (4.79) and equation (4.81) are solved using algorithm (4.83b).

$$P_1 = \begin{bmatrix} 1.1187 & 2.6472\text{e-}1 & -6.9279\text{e-}4 \\ 2.6472\text{e-}1 & 2.1096 & -6.5987\text{e-}3 \\ -6.9279\text{e-}4 & -6.5987\text{e-}3 & 2.1022 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 1.3242 & 2.5554\text{e-}1 & -1.2502\text{e-}3 \\ 2.5554\text{e-}1 & 1.6418 & -6.1620\text{e-}3 \\ -1.2502\text{e-}3 & -6.1620\text{e-}3 & 3.1012 \end{bmatrix},$$

$$\eta_{11} = \begin{bmatrix} 6.6542\text{e-}2 & 1.7460\text{e-}1 & 1.4531\text{e-}1 \\ 2.2761\text{e-}1 & 2.2356 & 3.3869\text{e-}1 \\ 1.8710\text{e-}2 & -6.5111\text{e-}1 & 1.4926\text{e-}1 \end{bmatrix},$$

$$\eta_{22} = \begin{bmatrix} 1.7417\text{e-}1 & 1.3506 & 1.4480 \\ -3.6335\text{e-}2 & -2.4805 & -1.0267 \\ 1.1609 & 8.1400 & 4.4421 \end{bmatrix}.$$

Algorithm (4.83) converges to the required solution with an accuracy of 1.0e-12 order after 28 and 8 iterations respectively. Furthermore, the incentive Stackelberg strategy (4.69) that

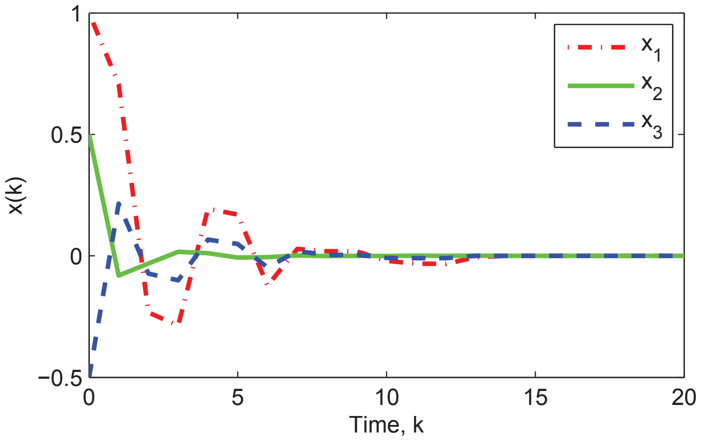


Fig. 4.6: Trajectory of the state.

will be announced by the leader can be calculated as,

$$u_{0i}^*(k) = \eta_{0i}x(k) + \eta_{ii}u_i^*(k), \tag{4.84}$$

where

$$\eta_{01} = \begin{bmatrix} -7.3449\text{e-}3 & -6.7328\text{e-}2 & 1.7155\text{e-}2 \\ 8.7881\text{e-}3 & -2.4716\text{e-}2 & -1.1732\text{e-}2 \\ -1.1877\text{e-}1 & -2.5325\text{e-}1 & 7.7147\text{e-}4 \end{bmatrix},$$

$$\eta_{02} = \begin{bmatrix} -2.0608\text{e-}2 & -7.8234\text{e-}2 & -3.5325\text{e-}3 \\ -7.9949\text{e-}2 & -1.4682\text{e-}1 & 1.5815\text{e-}2 \\ -9.9203\text{e-}2 & -2.0499\text{e-}1 & -5.7203\text{e-}2 \end{bmatrix}.$$

In fact, after announcing this incentive, the followers' strategy can be computed by applying the standard LQ theory.

$$u_i^*(k) = -[\tilde{R}_i + \tilde{B}_i^T P_i \tilde{B}_i]^{-1}[\tilde{B}_i^T P_i \tilde{A}_i + \eta_{ii}^T R_{0ii}\eta_{0i}]x(k), \tag{4.85}$$

which implies

$$u_1^*(k) = \begin{bmatrix} -1.3440\text{e-}2 & -2.8877\text{e-}2 & 1.3626\text{e-}3 \\ -1.1534\text{e-}2 & 4.2314\text{e-}3 & 4.9096\text{e-}3 \\ -3.7172\text{e-}2 & -8.2997\text{e-}2 & 2.7749\text{e-}3 \end{bmatrix} x(k),$$

$$u_2^*(k) = \begin{bmatrix} -1.7687\text{e-}2 & -3.1838\text{e-}2 & 6.5192\text{e-}4 \\ 1.9603\text{e-}2 & 4.4974\text{e-}2 & 4.9832\text{e-}3 \\ -1.4987\text{e-}2 & -3.8356\text{e-}2 & 3.7607\text{e-}3 \end{bmatrix} x(k).$$

Indeed, it can be observed that this matrix gain is equal to $\tilde{K}_i$. Namely, it can be confirmed that the followers adopt the team-optimal solution with the H_∞ constraint eventually.

Second, considering a finite time interval, the evaluation of $P(k)$ and $W(k)$ are given in Fig. 4.2, which clearly show the convergence of the solution of SMVDEs (4.56a) and (4.56c). In a similar manner, convergences of $P_i(k)$ corresponding to each i-th follower, as referred to SBDRE (4.64), are shown in Fig. 4.3. Moreover, by solving (4.66), we present the convergence of $\eta_{11}(k)$ and $\eta_{22}(k)$ in Fig. 4.4 and by (4.59), we present $\eta_{01}(k)$ and $\eta_{02}(k)$ in Fig. 4.5. It should be noted that for $(P,\ W)$, $(P_1,\ P_2)$, $(\eta_{11},\ \eta_{22})$ and $(\eta_{01},\ \eta_{02})$, the authenticity of results in the infinite-horizon case can be verified easily by comparing with Figures 4.2, 4.3, 4.4 and 4.5, respectively while we consider the time parameter, $k \to \infty$, in the finite-horizon case.

Fig. 4.6 shows the response of the system with a state trajectory. It should be noted that $w(k)$ is chosen as the stochastic-process-dependent random disturbance such that $\mathbb{E}[w(k)] = 0$ and $|w(.)| \leq 30$. Furthermore, Fig. 4.6 represents that the state variables $x(k)$ can stabilize the given system, which implies that the proposed method is very useful and reliable.

4.6.2 A simple practical example

In order to demonstrate the efficiency of our proposed strategies, an *R–L–C* electrical circuit that can be represented stochastic system in Fig. 4.7 is considered. In this network, R_i, r_i, $i = 1,\ 2$, R and L are the resistances and the inductance, respectively. The capacitances are denoted by C_i, $i = 1,\ 2$. Moreover, $E_{01}(t) := u_{01}(t)$, $E_{02}(t) := u_{02}(t)$ and $E_1(t) := u_1(t)$, $E_2(t) := u_2(t)$ denote the applied voltages, that is, the control inputs as the leaders and the followers, respectively. $i(t)$ denote the electric current in the inductance L.

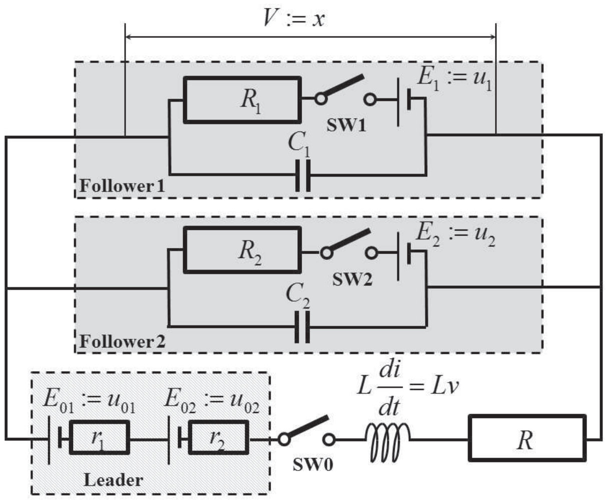


Fig. 4.7: Circuit diagram.

According to Fig. 4.7, for the Follower 1, current through the resistor:

$$i_{11} = \frac{V - E_1}{R_1}, \tag{4.86}$$

and current into the capacitor:

$$i_{12} = C_1 \frac{dV}{dt}. \tag{4.87}$$

Let us consider,

$$i_1 = i_{11} + i_{12}, \tag{4.88}$$

which implies that the current trough the Follower 1 is,

$$i_1 = \frac{V - E_1}{R_1} + C_1 \frac{dV}{dt}, \tag{4.89}$$

Similarly, current trough the Follower 2 is,

$$i_2 = \frac{V - E_2}{R_2} + C_2 \frac{dV}{dt}, \tag{4.90}$$

The voltage drop across the inductor is $L\frac{di}{dt}$. If we consider the total current flow through the circuit is,

$$i = i_1 + i_2, \tag{4.91}$$

we can get the following differential equation:

$$\frac{E_{01} + E_{02} - L\frac{di}{dt} - V}{R + r_1 + r_2} = (C_1 + C_2)\frac{dV}{dt} + \frac{V - E_1}{R_1} + \frac{V - E_2}{R_2}, \tag{4.92}$$

or,

$$\begin{aligned} \frac{dV}{dt} =& \frac{1}{C_1 + C_2}\Bigg[-\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R + r_1 + r_2} \right) V \\ & + \frac{E_{01} + E_{02}}{R + r_1 + r_2} + \frac{E_1}{R_1} + \frac{E_2}{R_2} - \frac{L\frac{di}{dt}}{R + r_1 + r_2,} \Bigg]. \end{aligned} \tag{4.93}$$

For this system, let us consider $di(t)/dt = v(t)$ as an external disturbance and the voltage drop across the circuit, $V := x$, as a state with initial voltage $x(0) = x_0$. It should be noted that, in any electronic device, thermal noise is unavoidable at non-zero temperatures. This means the system can be represented as a *stochastic differential equation* (SDE) with a random noise term. If this noise is treated as a real-valued state-dependent Wiener process

$w(t)$ with coefficient A_p, then the stochastic system can be written, in simplified notation, as

$$dx(t) = \left[\bar{A}x(t) + \sum_{i=1}^{2}[\bar{B}_{0i}u_{0i}(t) + \bar{B}_i u_i(t)] + \bar{D}v(t)\right] dt + \bar{A}_p x(t) dw(t), \quad x(0) = x_0. \quad (4.94)$$

where

$$\begin{aligned}
\bar{A} &:= -\frac{1}{C_T}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_S}\right), \\
\bar{B}_{01} = \bar{B}_{02} &:= \frac{1}{C_T R_S}, \\
\bar{B}_1 &:= \frac{1}{C_T R_1}, \\
\bar{B}_2 &:= \frac{1}{C_T R_2}, \\
\bar{D} &:= -\frac{L}{C_T R_S}, \\
\bar{A}_p &:= 0.1\bar{A}, \\
C_T &:= C_1 + C_2, \\
R_S &:= R + r_1 + r_2.
\end{aligned}$$

It should be noted that the system noise has been added to the deterministic system by describing it stochastically in the SDE (4.94). It is assumed that 1% of the magnitude of the state coefficient can be represented by a Wiener process based on stochastic perturbations.

Suppose that we wish to solve this SDE (4.94) for some time interval $[0,\ T]$. By the Euler-Maruyama approximation, the stochastic continuous-time system (4.94) can be transformed into a stochastic discrete-time system with T_f equal sub-intervals of width $\Delta t = T/T_f$, as follows:

$$\begin{cases}
x(k+1) = A(k)x(k) + \sum_{j=1}^{2}\left[B_{0j}(k)u_{0j}(k) + B_j(k)u_j(k)\right] \\
\qquad\qquad +D(k)v(k) + A_p x(k)N(0,1),\ x(0) = x_0, \\
z(k) = \begin{bmatrix} x(k) & u_{01}(k) & u_{02}(k) & u_1(k) & u_2(k) \end{bmatrix}^T,
\end{cases} \quad (4.95)$$

where

$$\begin{aligned}
A &:= 1 + \bar{A}\Delta t, \\
B_{01} &:= \bar{B}_{01}\Delta t, \\
B_{02} &:= \bar{B}_{02}\Delta t, \\
B_1 &:= \bar{B}_1\Delta t,
\end{aligned}$$

$$B_2 := \bar{B}_2 \Delta t,$$
$$D := \bar{D} \Delta t,$$
$$A_p := \bar{A}_p \sqrt{\Delta t},$$

and $N(0,1)$ denotes a normally distributed random variable with zero mean and unit variance.

In this problem, it is assumed that the leader will control the voltage sources in such a way that the team-optimal solution will be achieved, attenuating the external disturbance under the H_∞ constraint. On the contrary, the followers will simultaneously optimize their own costs using Nash equilibrium strategies, with respect to the leader's incentive Stackelberg strategy (4.46).

In order to solve this problem numerically, the simulation data are assigned to the parameters as follows:

$$R_1 = 2\ \text{M}\Omega,\ \ R_2 = 3\ \text{M}\Omega,\ \ R = 600\ \Omega,\ \ r_1 = 2\ \Omega,\ \ r_2 = 3\ \Omega,$$
$$C_1 = 1\ \mu\text{F},\ C_2 = 2\mu\text{F},\ L = 0.01\ \text{H},\ x_0 = 5\ \text{V},\ \Delta t = 1\ \text{ms}.$$

The weight matrices of the cost functionals of the leader and followers can be defined as in

$$R_{001} = 2,\ R_{002} = 4,\ R_{01} = 3,\ R_{02} = 2,\ R_{011} = 4,\ R_{022} = 4,$$
$$R_{11} = 3,\ R_{22} = 2,\ Q_0 = 1,\ Q_1 = 2,\ Q_2 = 4.$$

Now, we choose as $\gamma = 3$ to design the incentive Stackelberg strategy-set.

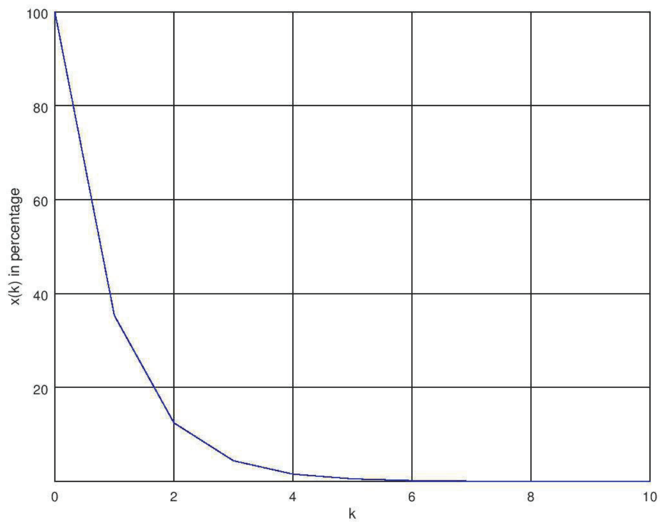


Fig. 4.8: Percentage of voltage discharging.

The H_∞-constrained team-optimal state-dependent strategies of the voltage sources can be determined to be

$$\begin{aligned}
u_{01}(k) &= -1.1563\text{e-}1x(k),\\
u_{02}(k) &= -5.7815\text{e-}2x(k),\\
u_{1}(k) &= -2.3319\text{e-}5x(k),\\
u_{2}(k) &= -3.4978\text{e-}5x(k).
\end{aligned}$$

In addition, the incentive Stackelberg strategies announced by the leader can be found to be

$$\begin{aligned}
u_{01}(k) &= -1.1563\text{e-}1x(k) + 1.1464\text{e-}2u_{1}(k),\\
u_{02}(k) &= -5.7815\text{e-}2x(k) - 3.0250\text{e-}4u_{2}(k).
\end{aligned}$$

Finally, Fig. 5 shows the exponential decay of the capacitors' voltage over time, calculated using our current method.

Remark 4.4. *As we know, LQ is a control scheme that provides optimal performance for certain given quadratic objective functions. However, it cannot be used for hierarchical decision making and optimal collaboration of multiple decision making problems. The results of our proposed method show that by applying the incentive Stackelberg strategy, the leader can impose his control over the followers.*

In many real-world control problems, most physical systems and processes contain undetected uncertainties in the deterministic exogenous disturbances. In contrast to stochastic perturbations due to the Wiener process, they introduce serious difficulties in the control and design of the system. Moreover, engineering practices not only reduce the impact of disturbances, but also minimize the desired objective in the presence of disturbance. Considering all of the above possible constraints, we propose such a solution to the problem reflecting the real-life system in more general.

For example, [Luo et al. (2016)] investigates the problem of stimulating users to participate in mobile crowdsourcing applications through personal devices such as smartphones/tablets/laptops. However, motivating enough users to provide their personal device resources to achieve the good quality of service is a challenge. To solve this problem, the authors propose an incentive framework based on Stackelberg game to simulate the interaction between the server and users.

4.7 Conclusion

This chapter investigates the incentive Stackelberg game for discrete-time stochastic systems. The motivation is to choose the incentive Stackelberg game is an engineering application of a packet switch that works in the loop structure [Saksena and Cruz (1985)]. The above problem comes from a static game. However, this chapter only studies dynamic games. This chapter involves one leader and multiple followers incentive Starkberg game. For this game, incentive Stackelberg strategy is the idea in which leader can implement his team-optimal solution in a Stackelberg game. In the followers' group, players are supposed to be non-cooperative, and Nash equilibrium is investigated. Unlike the previous chapter, this chapter examines the stochastic system.

The deterministic disturbances and their attenuation to stochastic systems under the H_∞ constraint is the main attraction of this chapter. Problems involving deterministic disturbance must be attenuated at a given target called disturbance attenuation level $\gamma > 0$. Surprisingly, the concept of solving the disturbance reduction problem under the H_∞ constraint seems like a Nash equilibrium between the disturbance input and the control input. In this game, an incentive structure is developed in such a way that leader achieve team-optimal solution attenuating the disturbance under H_∞ constraint. Simultaneously, followers achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance is considered. Results based on both finite- and infinite- time domains are shown in this chapter. The structure of the incentive Stackelberg game is the same in both finite- and infinite-horizon problems. The main focus of the infinite horizon situation is stochastic Lyapunov stability theory. Using stochastic Lyapunov stability theory, several lemmas have been proved. A computational algorithm based on Lyapunov iterations is developed to obtain the incentive Stackelberg strategy-set.

This chapter studies the most common linear quadratic (LQ) optimal control in the game problems. In order to solve the LQ problem, discrete-time stochastic dynamic programming (SDP) is deeply studied. Several basic lemmas are completely proved and useful for this chapter. The solution sets for incentive Stackelberg strategy are found by solving a set of stochastic backward difference Riccati equations (SBDREs) in the finite-horizon case. On the other hand, it is shown that the results of the infinite-horizon case are found by solving a set of stochastic algebraic Riccati equations (SAREs). In order to ensure the stability of the system, the state trajectory figure is presented. To demonstrate the effectiveness of the proposed method, an academic example and a practical example are presented. However, this chapter only investigates one leader game problem, which leads many leaders to further study.

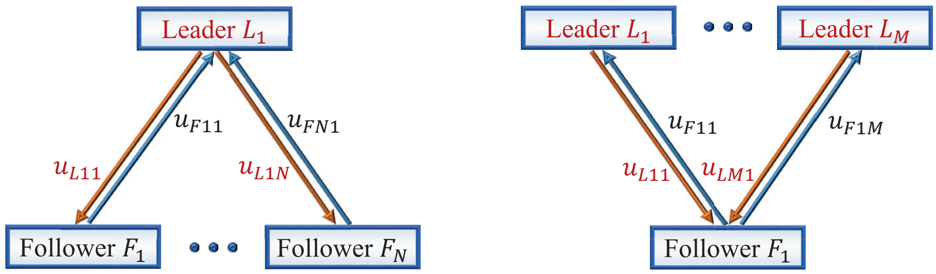
Chapter 5

Infinite-Horizon Multi-Leader-Follower Incentive Stackelberg Games for Linear Stochastic Systems with H_∞ Constraint

This chapter is based on a previously published article [Ahmed et al. (2017b)].

5.1 Introduction

Through the last four decades, incentive Stackelberg games are the growing interest in research (see, e.g., [Basar and Olsder (1980), Ho et al. (1982), Ishida and Shimemura (1983), Zheng et al. (1984), Mizukami and Wu (1987), Mizukami and Wu (1988)], and references therein). In a Stackelberg games, when the leader's strategy induces the follower's decision such that the leader's prespecified optimal solution or equilibrium (e.g., team-optimul, Nash equilibrium) can be achieved - called the incentive Stackelberg strategy. In [Mizukami and Wu (1987)] and [Mizukami and Wu (1988)], incentive Stackelberg strategies were derived for LQ differential games, where the two leaders and one follower to the first paper and one leader and two followers to the second paper were considered. Recently, in [Mukaidani and Xu (2018)], one leader with multiple followers was considered for stochastic linear system with H_∞ constraint. Moreover, the similar structure also applied for deterministic discrete-time case in [Ahmed and Mukaidani (2016)]. On the other hand, multiple leaders and one follower incentive Stackelberg games were investigated in [Mukaidani (2016)] for infinite-horizon stochastic linear system. However, in practical engineering or social systems, it is generally assumed that there exists multiple leaders and a large number of followers, with the leaders being multiple players. Thus, it is natural to consider possible decision patterns among the multiple leaders and followers.



(a) One leader N followers

(b) M leaders one follower

Fig. 5.1: Two-level multi-leader-follower hierarchy.

In this chapter, an infinite-horizon continuous-time incentive Stackelberg game for a class of linear stochastic systems governed by the Itô's differential equation involving multi-leader-follower is investigated. In this game, a deterministic exogenous disturbance is also observed, which is attenuating under H_∞ constraint. Recently, in [Mukaidani et al. (2017b)], incentive Stackelberg strategy with multiple leaders and multiple followers are considered for the stochastic Markov jumping control problem. To the best of our knowledge, there have been no studies on exogenous disturbance and its attenuation under H_∞ constraint, which makes the game more complicated to deal with. Hence, our current research is not simply a trivial extension of existing studies. Moreover, unlike the deterministic system [Mizukami and Wu (1987)], [Mizukami and Wu (1988)], stochastic incentive Stackelberg games involving state-dependent noise with deterministic external disturbance are studied for the first time. It should be noted that we only discuss two-level hierarchical games with M leaders and N followers.

To illustrate such a multi-players hierarchical game, we consider that M leaders and N followers are belonging to two groups. L_1, L_2, ... ,L_M are in the leader's group and F_1, F_2, ... ,F_N are in the follower's group. Note that, when $M = 1$ the structure of the game is the same as the structure used in [Mukaidani and Xu (2018)] depicted in Fig. 5.1(a). On the other hand, if $N = 1$ the structure of the game will be same as [Mukaidani (2016)] depicted in Fig. 5.1(b). Furthermore, tuning the value of M and N, we can form any convenient hierarchical game. That's why, we have termed it as a generalized structure. We establish the following patterns for this game,

- Each leader achieves Nash equilibrium solution attenuating the exterior disturbance under H_∞ constraint.
- Each leader declares incentive Stackelberg strategies for each follower, individually.

- Each follower will adopt the Nash equilibrium strategy regarding leader's incentive strategies declared in advance.
- Leaders and followers act non-cooperatively in their own group.

In our work, the conditions for the leader's Nash equilibrium strategy under the H_∞ constraint are derived based on the existing results for the infinite-horizon stochastic H_2/H_∞ control problem in [Chen and Zhang (2004)]. It is shown that the strategy-set can be found by solving some *cross-coupled stochastic algebraic Riccati equations* (CCSAREs) and *matrix algebraic equations* (MAEs). The leader's and the follower's strategies are established in such a way that simultaneously leaders achieve Nash equilibrium strategies attenuating the external disturbance under H_∞ constraint ensuring the follower's Nash equilibrium strategy while worst-case disturbance is considered. However, the derivation of the CCSAREs seems not so easy; the solution can be found easily by Lyapunov based iterations. A simple numerical example is provided to illustrate our findings.
Notation: The notations used in this chapter are fairly standard. I_n denotes the $n \times n$ identity matrix. $\mathbf{col}[\cdot]$ denotes a column vector. $\mathbf{diag}[\cdot]$ denotes a diagonal matrix. $\mathbf{blockdiag}[\cdot]$ denotes a block diagonal matrix. $\|\cdot\|$ denotes the Euclidean norm. $\mathbb{E}[\cdot]$ denotes the expectation operator. $\mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^\ell)$ denotes the space of nonanticipative stochastic processes. Finally, for an N-tuple

$$\gamma = (\gamma_1,\ \ldots\ ,\gamma_N) \in \Gamma_1 \times \ldots \times \Gamma_N,$$

and for given sets Γ_i, we write

$$\gamma^*_{-i}(\alpha) := (\gamma^*_1,\ \ldots\ ,\gamma^*_{i-1}, \alpha, \gamma^*_{i+1},\ \ldots\ ,\gamma^*_N),$$

with $\alpha \in \Gamma_i$.

5.2 Definitions and preliminaries

In this section, we will introduce stochastic H_2/H_∞ control problem and exact observability.

Consider the following stochastic linear system:

$$dx(t) = [Ax(t) + Bu(t) + Ev(t)]\,dt + A_p(t)x(t)dw, \qquad x(0) = x^0 \tag{5.1a}$$
$$z(t) = \mathbf{col}\begin{bmatrix} Cx(t) & Du(t) \end{bmatrix}, \quad D^T D = I, \tag{5.1b}$$

where, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ is the control input, $v(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v})$ is the deterministic disturbance, $w(t) \in \mathbb{R}$ is a one-dimensional wiener process and $z(t) \in \mathbb{R}^{n_z}$ is the controlled output. The infinite-horizon stochastic H_2/H_∞ control problem can be stated as follows [Chen and Zhang (2004)],

Definition 5.1. *Given the disturbance attenuation level $\gamma > 0$, to find $u^*(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ and a worst-case disturbance $v^*(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v})$ such that*
(i) when the optimal state feedback control $u^*(t) = K^* x(t)$ is applied, then

$$\|\mathbb{L}\|_\infty := \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\|z\|}{\|v\|} = \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\sqrt{\mathbb{E}\left[\int_0^\infty \{x^T C^T C x + u^T u\}\, dt\right]}}{\sqrt{\mathbb{E}\left[\int_0^\infty v^T v\, dt\right]}} < \gamma, \tag{5.2}$$

where $\mathbb{L}(v) = \begin{bmatrix} Cx(t,u^*,v,0) \\ Du^* \end{bmatrix}$ *is called the perturbed operator of* (5.1).

(ii) $u^*(t)$ stabilizes the system (5.1) internally, i.e.

$$\lim_{t\to\infty} \mathbb{E}\|x(t,u^*,0,x^0)\|^2 = 0. \tag{5.3}$$

(iii) when the worst-case disturbance $v^*(t) = K^*_\gamma x(t)$ is applied, $u^*(x)$ minimizes the output cost

$$J_u(u,v^*) = \|z\|_2^2 = \mathbb{E}\left[\int_0^\infty \left\{x^T C^T C x + u^T u\right\} dt\right]. \tag{5.4}$$

Here, worst-case disturbance means

$$v^*(t) = \underset{v}{\text{argmin}}\ J_v(u^*,v), \tag{5.5}$$

with

$$J_v(u^*,v) = \mathbb{E}\left[\int_0^\infty (\gamma^2\|v\|^2 - \|z\|^2) dt\right]. \tag{5.6}$$

If the above mentioned $(u^*(t),\ v^*(t))$ exists, then we say that the infinite-horizon H_2/H_∞ control admits a pair of solutions.

Remark 5.1. *The infinite-horizon stochastic H_2/H_∞ control strategy pair $(u^*(t),\ v^*(t))$ is associated with the Nash equilibrium strategies.*

Lemma 5.1. *[Chen and Zhang (2004)] For the system* (5.1)*, suppose the CCSAREs:*

$$XA_l + A_l^T X + A_p^T X A_p + XS_l X + C^T C = 0, \tag{5.7a}$$
$$YA_l + A_l^T Y + A_p^T Y A_p - \gamma^{-2} YTY + XS_l X + C^T C = 0, \tag{5.7b}$$

with

$$A_l := A - S_l X + \gamma^{-2} T Y, \quad S_l := BB^T, \quad T := EE^T,$$

have solutions $X > 0$, $Y > 0$. *If* $[A,\ A_p|C]$ *and* $[A - \gamma^{-2} DD^T Y,\ A_p|C]$ *are exactly observable, then the stochastic* H_2/H_∞ *control problem admits a solution set:*

$$u^*(t) = K^* x(t) = -B^T X x(t), \tag{5.8}$$

$$v^*(t) = K_\gamma^* x(t) = \gamma^{-2} E^T Y x(t). \tag{5.9}$$

Proof. This lemma has been proved earlier as Theorem 2.8 in Chapter 2. □

5.3 Problem formulation

Consider a linear stochastic system governed by the Itô differential equation defined by

$$\begin{aligned} dx(t) = \Bigg[& Ax(t) + \sum_{i=1}^{M} [B_{Li1} u_{Li1}(t) + \ \dots \ + B_{LiN} u_{LiM}(t)] \\ & + \sum_{j=1}^{N} [B_{Fj1} u_{Fj1}(t) + \ \dots \ + B_{FjM} u_{FjM}(t)] + Dv(t) \Bigg] dt + A_p x(t) dw(t), \ \ x(0) = x^0, \end{aligned} \tag{5.10a}$$

$$z(t) = \mathbf{col}\left[Cx(t) \quad u_{c1}(t) \ \dots \ u_{cM}(t) \right], \tag{5.10b}$$

$$u_{ci}(t) = \mathbf{col}\left[u_{Li1}(t) \ \dots \ u_{LiN}(t) \quad u_{F1i}(t) \ \dots \ u_{FNi}(t) \right], \tag{5.10c}$$

where $x(t) \in \mathbb{R}^n$ represents the state vector; $z(t) \in \mathbb{R}^{n_z}$ represents the controlled output; $u_{Lij}(t) \in \mathbb{R}^{m_{Lij}}$ represents the leader L_i's control input for the follower F_j, $i = 1, \ \dots \ ,M, \ j = 1, \ \dots \ ,N$; $u_{Fji}(t) \in \mathbb{R}^{m_{Fji}}$ represents the follower F_j's control input according to the leader L_i, $i = 1, \ \dots \ ,M, \ j = 1, \ \dots \ ,N$; $v(t) \in \mathbb{R}^{m_v}$ represents the exogenous disturbance signal; $w(t) \in \mathbb{R}$ represents a one-dimensional standard Wiener process defined in the filtered probability space $(\Omega,\ \mathscr{F},\ \mathscr{P},\ \mathscr{F}_t)$ with $\mathscr{F}_t = \sigma\{w(s) \ : \ 0 \le s \le t\}$ [Chen and Zhang (2004)]. Cost functionals of the leaders L_i, $i = 1, \ \dots \ ,M$, are given by

$$\begin{aligned} & J_{Li}(u_{Li1}, \ \dots \ ,u_{LiN},\ u_{F1i}, \ \dots \ ,u_{FNi},\ v) \\ & := \frac{1}{2} \mathbb{E}\left[\int_0^\infty \left\{ x^T(t) Q_{Li} x(t) + \sum_{j=1}^{N} \left[u_{Lij}^T(t) R_{Lij} u_{Lij}(t) + u_{Fji}^T(t) R_{LFji} u_{Fji}(t) \right] \right\} dt \right], \end{aligned} \tag{5.11}$$

where $Q_{Li} = Q_{Li}^T \ge 0$, $R_{Lij} = R_{Lij}^T > 0$, $R_{LFji} = R_{LFji}^T \ge 0$, $i = 1, \ \dots \ ,M, \ j = 1, \ \dots \ ,N$. Cost functionals of the followers F_i, $i = 1, \ \dots \ ,N$ are given by

$$J_{Fi}(u_{L1i}, \ \dots \ ,u_{LMi},\ u_{Fi1}, \ \dots \ ,u_{FiM},\ v)$$

$$:= \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)Q_{Fi}x(t) + \sum_{j=1}^{M}\left[u_{Lji}^T(t)R_{FLji}u_{Lji}(t) + u_{Fij}^T(t)R_{Fij}u_{Fij}(t)\right]\right\}dt\right], \tag{5.12}$$

where $Q_{Fj} = Q_{Fj}^T \geq 0$, $R_{Fij} = R_{Fij}^T > 0$ and $R_{FLji} = R_{FLji}^T \geq 0$, $i = 1, \ldots, N,\ j = 1, \ldots, M$. For a two-level incentive Stackelberg game, leaders announce the following incentive strategy to the followers in ahead of time:

$$u_{Lij}(t) = \Lambda_{ji}x(t) + \Xi_{ji}u_{Fji}(t), \quad i = 1, \ldots, M \quad j = 1, \ldots, N, \tag{5.13}$$

where the parameters Λ_{ji} and Ξ_{ji} are to be determined associated with the Nash equilibrium strategies $u_{Fji}(t)$ of the followers for $i, \ldots, M\ j = 1, \ldots, N$. In this game, leaders will achieve a Nash equilibrium solution attenuating the external disturbance with H_∞ constraint. The infinite-horizon multi-leader-follower incentive Stackelberg games for linear stochastic systems with H_∞ constraint can be formulated as follows.

For any disturbance attenuation level $\gamma > 0$, to find, if possible, the state feedback strategy $u_{Lij}^*(t) = K_{cij}x(t)$ and $u_{Fji}^*(t) = K_{Fji}x(t)$ such that
(i) the trajectory of the closed-loop system (5.10) satisfies the Nash equilibrium conditions (5.14a) of the leaders with H_∞ constraint condition (5.14b):

$$J_{Li}(u_{c1}^*, \ldots, u_{cM}^*, v^*) \leq J_{Li}(\gamma_{-i}^*(u_{ci})), v^*), \tag{5.14a}$$

$$0 \leq J_v(u_{c1}^*, \ldots, u_{cM}^*, v^*) \leq J_v(u_{c1}^*, \ldots, u_{cM}^*, v), \tag{5.14b}$$

where $i = 1, \ldots, M$,

$$J_v(u_{c1}, \ldots, u_{cM}, v) = \mathbb{E}\left[\int_0^\infty \left\{\gamma^2\|v(t)\|^2 - \|z(t)\|^2\right\}dt\right], \tag{5.15}$$

$$\|z(t)\|^2 = x^T(t)C^TCx(t) + \sum_{i=1}^{M} u_{ci}^T(t)u_{ci}(t), \tag{5.16}$$

$\forall\ v(t) \neq 0 \in \mathbb{R}^{m_v}$,
(ii) with a worst-case disturbance $v^*(t) \in \mathbb{R}^{m_v}$, follower's decision $u_{Fji}^*(t)) \in \mathbb{R}^{m_{Fji}}$; $i = 1, \ldots, M,\ j = 1, \ldots, N$ satisfies the following Nash equilibrium conditions:

$$J_{Fj}(u_{F1}^*, \ldots, u_{FN}^*, v^*) \leq J_{Fj}(\gamma_{-j}^*(\hat{u}_{Fj})), v^*), \tag{5.17}$$

where

$$\hat{u}_{Fj}(t) = \mathbf{col}\left[u_{Fj1}(t) \quad \ldots \quad u_{FjM}(t)\right],\ j = 1, \ldots, N.$$

It should be noted that $u_{Lij}(t)$ depend on $u_{Fji}(t)$ according to the incentive Stackelberg structures assumed in (5.13).

Remark 5.2. *If the inequality* (5.14b) *holds, we say that it satisfies the condition for H_∞ constraint,*

$$\|\mathbb{L}\|_\infty = \sup_{v\in\mathbb{R}^{m_v},\ v\neq 0,\ x^0=0} \frac{\|z\|_2}{\|v\|_2} < \gamma. \tag{5.18}$$

A set of decision $(u_{c1}^*(t),\ \dots\ ,u_{cM}^*(t), v^*)$ is said to constitute a Nash equilibrium solution for a two-level hierarchical game with H_∞ constraint.

Without loss of generality, some assumptions for the following decision process are made [Basar and Selbuz (1979), Mizukami and Wu (1987)].

Assumption 5.1. *Leader L_i announces his/her own strategy of the feedback pattern* (5.13) *ahead of time to the followers F_j as the incentive strategies.*

Assumption 5.2. *Followers F_j decide their optimal strategies once they know the strategy announced by L_i.*

Assumption 5.3. *L_i and F_j at the same level act non-cooperatively and they decide Nash equilibrium.*

5.4 Main results

Let us define the space of admissible strategies for the players L_{ij} by Γ_{Lij} and for the players F_{ji} by Γ_{Fji}, $i = 1,\ \dots\ ,M,\ j = 1,\ \dots\ ,N$. For each pair $(u_{Lij},\ u_{Fji}) \in \Gamma_{Lij} \times \Gamma_{Fji}$, it is supposed that the linear stochastic systems (5.10) has a unique solution on $0 \le t < \infty$ for all $x^0 \in \mathbb{R}^n$ and the values of J_{Li} and J_{Fj} are well defined.

5.4.1 Leader's Nash equilibrium strategy

First, the H_∞ constraint Nash equilibrium solutions $(u_{c1}^*(t),\ \dots\ ,u_{cM}^*(t), v^*)$ for the leaders are investigated. By composing the stochastic system (5.10a), the following centralized systems can be obtained.

$$dx(t) = \left[Ax(t) + \sum_{i=1}^{M} B_{ci}u_{ci}(t) + Dv(t)\right]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{5.19}$$

where

$$B_{ci} = \begin{bmatrix} B_{Li1} & \dots & B_{LiN} & B_{F1i} & \dots & B_{FNi} \end{bmatrix},\ i = 1,\ \dots\ ,M.$$

Furthermore, the cost functional (5.11) can be modified as

$$J_{Li}(u_{ci}(t)) = \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)Q_{Li}x(t) + u_{ci}^T(t)R_{ci}u_{ci}(t)\right\}dt\right], \quad i = 1,\ \dots\ ,M, \tag{5.20}$$

where

$$R_{ci} = \textbf{block diag}\begin{bmatrix} R_{Li1} & \dots & R_{LiN} & R_{LF1i} & \dots & R_{LFNi} \end{bmatrix}.$$

Therefore, for the Nash equilibrium solution with H_∞ constraint, the following result can be obtained from Lemma 5.1.

For the system (5.19), suppose the CCSAREs:

$$P_{ci}A_c + A_c^T P_{ci} + A_p^T P_{ci} A_p + P_{ci} S_{ci} P_{ci} + Q_{Li} = 0, \tag{5.21a}$$

$$WA_c + A_c^T W + A_p^T W A_p - \gamma^{-2} WTW + Q_L = 0, \tag{5.21b}$$

with

$$A_c := A - \sum_{i=1}^{M} S_{ci} P_{ci} + \gamma^{-2} TW, \quad S_{ci} := B_{ci} R_{ci}^{-1} B_{ci}^T, \quad T := DD^T, \quad Q_L = \sum_{i=1}^{M} P_{ci} S_{ci} P_{ci} + C^T C,$$

have solutions $P_{ci} > 0$, $W > 0$. If $[A,\ A_p|C]$ and $[A - \gamma^{-2} DD^T W,\ A_p|C]$ are exactly observable, then the stochastic H_∞ constraint problem admits a solution set:

$$\begin{bmatrix} u_{ci}^*(t) \\ v^*(t) \end{bmatrix} = \begin{bmatrix} K_{ci}^* \\ K_\gamma^* \end{bmatrix} x(t) = \begin{bmatrix} -R_{ci}^{-1} B_{ci}^T P_{ci} \\ \gamma^{-2} D^T W \end{bmatrix} x(t), \tag{5.22}$$

where

$$K_{ci}^* = \begin{bmatrix} K_{Li1}^{*T} & \dots & K_{LiN}^{*T} & K_{F1i}^{*T} & \dots & K_{FNi}^{*T} \end{bmatrix}^T,\ i = 1,\ \dots,M.$$

It should be noted that the relation between Λ_{ji} and Ξ_{ji}, $i = 1,\ \dots,M,\ j = 1,\ \dots,N$ can be derived from (5.13) as

$$\begin{aligned} \Lambda_{ji} &= K_{Lij}^* - \Xi_{ji} K_{Fji}^*, \\ &= -R_{Lij}^{-1} B_{Lij}^T P_{ci} + \Xi_{ji} R_{LFji}^{-1} B_{Fji}^T P_{ci}, \quad i = 1,\ \dots,M, \quad j = 1,\ \dots,N. \end{aligned} \tag{5.23}$$

So, the leader's incentive Stackelberg strategy for the followers can be determined by

$$\begin{aligned} u_{Lij}(t) &= \left[K_{Lij}^* - \Xi_{ji} K_{Fji}^*(t)\right] x(t) + \Xi_{ji} u_{Fji}(t), \\ &= K_{Lij}^* x(t) + \Xi_{ji}\left[u_{Fji} - K_{Fji}^*(t) x(t)\right] \quad i = 1,\ \dots,M, \quad j = 1,\ \dots,N. \end{aligned} \tag{5.24}$$

To determine Ξ_{ji}, $\ i = 1,\ \dots,M,\ j = 1,\ \dots,N$, which satisfy (5.13), let us consider the following optimization problem.

5.4.2 Follower's Nash equilibrium strategy

To establish the Nash equilibrium for the followers according to the leader's incentive Stackelberg strategy (5.13) and worst-case disturbance $v^*(t)$, we get the following system from the system (5.10a),

$$
\begin{aligned}
dx(t) =& \Bigg[Ax(t) + \sum_{i=1}^{N}[B_{Fi1}u_{Fi1}(t) + \ \dots \ + B_{FiM}u_{FiM}(t)] \\
&+ \sum_{j=1}^{M}[B_{Lj1}u_{Lj1}(t) + \ \dots \ + B_{LjN}u_{LjN}(t)] + Dv(t)\Bigg]dt + A_p x(t)dw(t), \ \ x(0) = x^0, \\
=& \Bigg[Ax(t) + \sum_{i=1}^{N}[B_{Fi1}u_{Fi1}(t) + \ \dots \ + B_{FiM}u_{FiM}(t)] \\
&+ \sum_{j=1}^{M}[B_{Lj1}(\Lambda_{1j}x(t) + \Xi_{1j}u_{F1j}(t)) + \ \dots \ + B_{LjN}(\Lambda_{Nj}x(t) + \Xi_{Nj}u_{FNj}(t))] \\
&+ \gamma^{-2}DD^T W x(t)\Bigg]dt + A_p x(t)dw(t), \quad x(0) = x^0, \\
=& \Bigg[\hat{A}x(t) + \sum_{i=1}^{N}\hat{B}_{Fi}\hat{u}_{Fi}(t)\Bigg]dt + A_p x(t)dw(t), \quad x(0) = x^0,
\end{aligned} \tag{5.25}
$$

where

$$
\begin{aligned}
&\hat{A} := A + \sum_{j=1}^{M}[B_{Lj1}\Lambda_{1j} + \ \dots \ + B_{LjN}\Lambda_{Nj}] + \gamma^{-2}DD^T W, \\
&\hat{u}_{Fi} := \mathbf{c}ol\begin{bmatrix} u_{Fi1} & \dots & u_{FiM}\end{bmatrix}, \\
&\hat{B}_{Fi} := \begin{bmatrix} \bar{B}_{Fi1} & \dots & \bar{B}_{FiM}\end{bmatrix}, \\
&\bar{B}_{Fij} := B_{Fij} + B_{Lji}\Xi_{ij}, \quad i = 1, \ \dots \ , N, \ j = 1, \ \dots \ , M.
\end{aligned}
$$

The cost functional of i-th follower can written as

$$
\begin{aligned}
&J_{Fi}(u_{L1i}, \ \dots \ , u_{LMi}, \ u_{Fi1}, \ \dots \ , u_{FiM}, \ v) \\
&\quad = \frac{1}{2}\mathbb{E}\Bigg[\int_0^\infty \Bigg\{x^T(t)Q_{Fi}x(t) + \sum_{j=1}^{M}\Big[[\Lambda_{ij}x(t) + \Xi_{ij}u_{Fij}(t)]^T R_{FLji}[\Lambda_{ij}x(t) + \Xi_{ij}u_{Fij}(t)] \\
&\qquad + u_{Fij}^T(t)R_{Fij}u_{Fij}(t)\Big]\Bigg\}dt\Bigg], \quad i = 1, \ \dots \ , N,
\end{aligned} \tag{5.26}
$$

or, equivalently,

$$
J_{Fi}(\hat{u}_{Fi}) = \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)\hat{Q}_{Fi}x(t) + 2x^T(t)\hat{S}_{Fi}\hat{u}_{Fi}(t) + \hat{u}_{Fi}^T(t)\hat{R}_{Fi}\hat{u}_{Fi}(t)\right\}dt\right], \tag{5.27}
$$

where

$$
\begin{aligned}
&\hat{Q}_{Fi} := Q_{Fi} + \sum_{j=1}^{M} \Lambda_{ij}^T R_{FLji} \Lambda_{ij}, \\
&\hat{R}_{Fi} := \textbf{block diag} \begin{bmatrix} \bar{R}_{Fi1} & \dots & \bar{R}_{FiM} \end{bmatrix}, \\
&\bar{R}_{Fij} := R_{Fij} + \Xi_{ij}^T R_{FLji} \Xi_{ij}, \\
&\hat{S}_{Fi} := \begin{bmatrix} \bar{S}_{Fi1} & \dots & \bar{S}_{FiM} \end{bmatrix}, \\
&\bar{S}_{Fij} := \Lambda_{ij}^T R_{FLji} \Xi_{ij}, \quad i = 1, \dots, N, \ j = 1, \dots, M.
\end{aligned}
$$

It should be noted that there exists a cross-coupling term $2x^T(t)\hat{S}_{Fi}\hat{u}_{Fi}(t)$ in the cost functional (5.27). By using the technique similar to the one used in the stochastic optimal control problem [Chen and Zhang (2004)], follower's Nash strategy $u_{Fi}(t) = \hat{u}_{Fi}(t)$ can be obtained.

$$
\hat{u}_{Fi}(t) = -\hat{R}_{Fi}^{-1}(P_{Fi}\hat{B}_{Fi} + \hat{S}_{Fi})^T x(t) = K_{Fi}^{\dagger} x(t) = \begin{bmatrix} K_{Fi1}^{\dagger} x(t) \\ \vdots \\ K_{FiM}^{\dagger} x(t) \end{bmatrix} \tag{5.28}
$$

where

$$
K_{Fi}^{\dagger} = -\hat{R}_{Fi}^{-1}(P_{Fi}\hat{B}_{Fi} + \hat{S}_{Fi})^T, \tag{5.29}
$$

and $P_{Fi}, \ i = 1, \dots N$ are the symmetric non-negative solution of the following CCSAREs:

$$
P_{Fi}A_{Fi} + A_{Fi}^T P_{Fi} + A_p^T P_{Fi} A_p - (P_{Fi}\hat{B}_{Fi} + \hat{S}_{Fi})\hat{R}_{Fi}^{-1}(P_{Fi}\hat{B}_{Fi} + \hat{S}_{Fi})^T + \hat{Q}_{Fi} = 0, \tag{5.30}
$$

where

$$
A_{Fi} = \hat{A} + \sum_{k=1,\ k \neq i}^{N} \hat{B}_{Fk} K_{Fk}^{\dagger}, \quad i = 1, \dots, N. \tag{5.31}
$$

Furthermore, from (5.22) we can find

$$
K_{cFi}^{*} = \begin{bmatrix} K_{Fi1}^{*} \\ \vdots \\ K_{FiM}^{*} \end{bmatrix} = \begin{bmatrix} -R_{LFi1}^{-1} B_{Fi1}^T P_{c1} \\ \vdots \\ -R_{LFiM}^{-1} B_{FiM}^T P_{cM} \end{bmatrix} \quad i = 1, \dots, N. \tag{5.32}
$$

Furthermore, Ξ_{ij} satisfies the equivalence relation $K_{cFi}^{*} \equiv K_{Fi}^{\dagger}$ can establish from (5.29) and (5.32) as follows:

$$
\begin{bmatrix} -R_{LFi1}^{-1} B_{Fi1}^T P_{c1} \\ \vdots \\ -R_{LFiM}^{-1} B_{FiM}^T P_{cM} \end{bmatrix} = -\hat{R}_{Fi}^{-1}(P_{Fi}\hat{B}_{Fi} + \hat{S}_{Fi})^T. \tag{5.33}
$$

Equivalently,

$$\begin{bmatrix} R_{Fi1}+\Xi_{i1}^T R_{FL1i}\Xi_{i1} & & \\ & \ddots & \\ & & R_{FiM}+\Xi_{iM}^T R_{FLMi}\Xi_{iM} \end{bmatrix} \begin{bmatrix} R_{LFi1}^{-1}B_{Fi1}^T P_{c1} \\ \vdots \\ R_{LFiM}^{-1}B_{FiM}^T P_{cM} \end{bmatrix}$$
$$= \begin{bmatrix} \left(B_{Fi1}^T+\Xi_{i1}^T B_{L1i}^T\right)P_{Fi} \\ \vdots \\ \left(B_{FiM}^T+\Xi_{iM}^T B_{LMi}^T\right)P_{Fi} \end{bmatrix} + \begin{bmatrix} \Xi_{i1}^T R_{FL1i}\Lambda_{i1} \\ \vdots \\ \Xi_{iM}^T R_{FLMi}\Lambda_{iM} \end{bmatrix}. \tag{5.34}$$

Comparing each rows from both sides of (5.34), we get

$$[R_{Fij}+\Xi_{ij}^T R_{FLji}\Xi_{ij}]R_{LFij}^{-1}B_{Fij}^T P_{cj} = B_{Fij}^T P_{Fi}+\Xi_{ij}^T B_{Lji}^T P_{Fi}+\Xi_{ij}^T R_{FLji}\Lambda_{ij}. \tag{5.35}$$

By using relation (5.23) for Λ_{ij}, we get

$$\begin{aligned} &R_{Fij}R_{LFij}^{-1}B_{Fij}^T P_{cj}+\Xi_{ij}^T R_{FLji}\Xi_{ij}R_{LFij}^{-1}B_{Fij}^T P_{cj} = \\ &\quad B_{Fij}^T P_{Fi}+\Xi_{ij}^T B_{Lji}^T P_{Fi}-\Xi_{ij}^T R_{FLji}R_{Lji}^{-1}B_{Lji}^T P_{cj}+\Xi_{ij}^T R_{FLji}\Xi_{ij}R_{LFij}^{-1}B_{Fij}^T P_{cj}. \end{aligned} \tag{5.36}$$

Canceling the term $\Xi_{ij}^T R_{FLji}\Xi_{ij}R_{LFij}^{-1}B_{Fij}^T P_{cj}$ from both sides of (5.36) we get

$$R_{Fij}R_{LFij}^{-1}B_{Fij}^T P_{cj} = B_{Fij}^T P_{Fi}+\Xi_{ij}^T B_{Lji}^T P_{Fi}-\Xi_{ij}^T R_{FLji}R_{Lji}^{-1}B_{Lji}^T P_{cj}, \tag{5.37}$$

and after simplification, the following MAEs can be found:

$$\Xi_{ij}^T(B_{Lji}^T P_{Fi}-R_{FLji}R_{Lji}^{-1}B_{Lji}^T P_{cj}) = R_{Fij}R_{LFij}^{-1}B_{Fij}^T P_{ci}-B_{Fij}^T P_{Fi},\ i=1,\ \dots\ ,N,\ j=1,\ \dots\ ,M. \tag{5.38}$$

Remark 5.3. *It should be noted that the incentive parameter Ξ_{ij} can be uniquely determined if and only if $(B_{Lji}^T P_{Fi}-R_{FLji}R_{Lji}^{-1}B_{Lji}^T P_{cj})$ is non-singular.*

Theorem 5.1. *Suppose that the CCSAREs in* (5.21)*, CCSAREs* (5.30) *and the MAEs* (5.38) *have solutions. Then the strategy-set* (5.13) *under* (5.22) *and* (5.29) *constitutes the two-level incentive Stackelberg strategies with H_∞ constraint.*

5.5 Numerical example

In order to demonstrate the efficiency of our proposed strategies, a numerical example is investigated. Let us consider the following system matrices:

$$A = \begin{bmatrix} 0.92 & 0 \\ 1.23 & -2.9 \end{bmatrix}, \quad A_p = 0.1A,$$
$$B_{L11} = \begin{bmatrix} 0.13 & 0.20 \\ -0.55 & 0.81 \end{bmatrix}, \quad B_{L12} = \begin{bmatrix} 0.31 & 1.20 \\ -1.25 & 1.02 \end{bmatrix},$$

$$B_{L21} = \begin{bmatrix} 0.28 & 0.12 \\ 5.32 & 0 \end{bmatrix}, \quad B_{L22} = \begin{bmatrix} 0.12 & 0.56 \\ 1.0 & 0.32 \end{bmatrix},$$
$$B_{F11} = \begin{bmatrix} 0.15 & -0.11 \\ 0.55 & 1.32 \end{bmatrix}, \quad B_{F12} = \begin{bmatrix} 0.51 & 0.54 \\ 0.21 & 1.21 \end{bmatrix},$$
$$B_{F21} = \begin{bmatrix} 0.23 & -0.45 \\ 0.28 & 2.96 \end{bmatrix}, \quad B_{F22} = \begin{bmatrix} 0.21 & 0.21 \\ 2.11 & 1.86 \end{bmatrix},$$
$$D = \begin{bmatrix} 0.054 & 0.043 \\ 0.023 & 0.013 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \end{bmatrix},$$
$$Q_{L1} = \mathbf{diag}\begin{bmatrix} 1 & 2 \end{bmatrix}, \quad Q_{L2} = \mathbf{diag}\begin{bmatrix} 3.1 & 2.4 \end{bmatrix},$$
$$Q_{F1} = \mathbf{diag}\begin{bmatrix} 2.1 & 1.3 \end{bmatrix}, \quad Q_{F2} = \mathbf{diag}\begin{bmatrix} 1.3 & 4.2 \end{bmatrix},$$
$$R_{L11} = \mathbf{diag}\begin{bmatrix} 1.5 & 2.5 \end{bmatrix}, \quad R_{L12} = \mathbf{diag}\begin{bmatrix} 1.5 & 2.2 \end{bmatrix},$$
$$R_{L21} = \mathbf{diag}\begin{bmatrix} 2.1 & 1.3 \end{bmatrix}, \quad R_{L22} = \mathbf{diag}\begin{bmatrix} 2.1 & 2.2 \end{bmatrix},$$
$$R_{F11} = \mathbf{diag}\begin{bmatrix} 2.3 & 1.1 \end{bmatrix}, \quad R_{F12} = \mathbf{diag}\begin{bmatrix} 1.8 & 4.2 \end{bmatrix},$$
$$R_{F21} = \mathbf{diag}\begin{bmatrix} 3.1 & 2.1 \end{bmatrix}, \quad R_{F22} = \mathbf{diag}\begin{bmatrix} 3.1 & 1.2 \end{bmatrix},$$
$$R_{LF11} = \mathbf{diag}\begin{bmatrix} 1.5 & 2.5 \end{bmatrix}, \quad R_{LF12} = \mathbf{diag}\begin{bmatrix} 2.1 & 1.2 \end{bmatrix},$$
$$R_{LF21} = \mathbf{diag}\begin{bmatrix} 2.1 & 1.3 \end{bmatrix}, \quad R_{LF22} = \mathbf{diag}\begin{bmatrix} 1.5 & 2.4 \end{bmatrix},$$
$$R_{FL11} = \mathbf{diag}\begin{bmatrix} 2.3 & 1.1 \end{bmatrix}, \quad R_{FL12} = \mathbf{diag}\begin{bmatrix} 1.9 & 1.2 \end{bmatrix},$$
$$R_{FL21} = \mathbf{diag}\begin{bmatrix} 3.1 & 2.1 \end{bmatrix}, \quad R_{FL22} = \mathbf{diag}\begin{bmatrix} 1.4 & 6.2 \end{bmatrix},$$

We choose the disturbance attenuation level as $\gamma = 5$. First, the CCAREs (5.21) are solved. These solutions attain the H_∞-constrained Nash equilibrium solutions set (5.22) as given below:

$$P_{c1} = \begin{bmatrix} 7.0074\text{e-}1 & -7.3498\text{e-}2 \\ -7.3498\text{e-}2 & 1.3855\text{e-}1 \end{bmatrix},$$
$$P_{c2} = \begin{bmatrix} 2.6369\text{e+}0 & -8.4572\text{e-}3 \\ -8.4572\text{e-}3 & 1.9503\text{e-}1 \end{bmatrix},$$
$$W = \begin{bmatrix} 1.7829\text{e+}0 & 1.6902\text{e-}1 \\ 1.6902\text{e-}1 & 3.0249\text{e-}1 \end{bmatrix},$$
$$K_{c1} = \left[\begin{array}{cc} -8.7680\text{e-}2 & 5.7171\text{e-}2 \\ -3.2246\text{e-}2 & -3.9010\text{e-}2 \\ \hline -2.0607\text{e-}1 & 1.3065\text{e-}1 \\ -3.4815\text{e-}1 & -2.4146\text{e-}2 \\ \hline -4.3125\text{e-}2 & -4.3451\text{e-}2 \\ 6.9640\text{e-}2 & -7.6388\text{e-}2 \\ \hline -6.6948\text{e-}2 & -1.0423\text{e-}2 \\ 4.0991\text{e-}1 & -3.4091\text{e-}1 \end{array}\right],$$

$$K_{c2} = \left[\begin{array}{cc} -3.3016\text{e-}1 & -4.9294\text{e-}1 \\ -2.4341\text{e-}1 & 7.8066\text{e-}4 \\ \hline -1.4665\text{e-}1 & -9.2387\text{e-}2 \\ -6.6999\text{e-}1 & -2.6215\text{e-}2 \\ \hline -6.3955\text{e-}1 & -1.7449\text{e-}2 \\ -1.1781\text{e+}0 & -1.9285\text{e-}1 \\ \hline -3.5727\text{e-}1 & -2.7315\text{e-}1 \\ -2.2418\text{e-}1 & -1.5041\text{e-}1 \end{array}\right],$$

$$K_{\gamma} = \begin{bmatrix} 4.0065\text{e-}3 & 6.4338\text{e-}4 \\ 3.1544\text{e-}3 & 4.4801\text{e-}4 \end{bmatrix},$$

Second, the CCAREs (5.30) and the MAEs (5.38) are solved as follows:

$$P_{F1} = \begin{bmatrix} 3.2114\text{e+}0 & 5.0554\text{e-}2 \\ 5.0554\text{e-}2 & 1.3704\text{e-}1 \end{bmatrix},$$

$$P_{F2} = \begin{bmatrix} 2.0458\text{e+}0 & -1.0578\text{e-}1 \\ -1.0578\text{e-}1 & 2.9737\text{e-}1 \end{bmatrix},$$

$$\Xi_{11} = \begin{bmatrix} 1.6632\text{e+}0 & -2.9164\text{e+}0 \\ -1.1162\text{e+}0 & 1.1706\text{e+}0 \end{bmatrix},$$

$$\Xi_{12} = \begin{bmatrix} 6.9087\text{e-}2 & -1.0437\text{e+}0 \\ 4.0327\text{e+}0 & -2.6264\text{e+}1 \end{bmatrix},$$

$$\Xi_{21} = \begin{bmatrix} 4.7540\text{e-}2 & 1.3001\text{e+}0 \\ -1.3020\text{e-}1 & -5.9341\text{e-}2 \end{bmatrix},$$

$$\Xi_{22} = \begin{bmatrix} 1.2908\text{e+}0 & -2.2264\text{e+}0 \\ -3.2405\text{e-}1 & 3.6127\text{e-}2 \end{bmatrix}.$$

Finally, the remaining parameter matrix (5.23) can be determined as,

$$\Lambda 11 = \begin{bmatrix} 1.8714\text{e-}1 & -9.3337\text{e-}2 \\ -1.6190\text{e-}1 & 1.9065\text{e-}3 \end{bmatrix},$$

$$\Lambda_{12} = \begin{bmatrix} -7.3580\text{e-}1 & 5.7434\text{e-}1 \\ -3.3254\text{e-}1 & -4.5733\text{e-}2 \end{bmatrix},$$

$$\Lambda_{21} = \begin{bmatrix} -1.5156\text{e+}0 & -6.9301\text{e-}1 \\ -2.8605\text{e+}1 & -4.9937\text{e+}0 \end{bmatrix},$$

$$\Lambda_{22} = \begin{bmatrix} -1.8459\text{e-}1 & -7.4664\text{e-}2 \\ -7.7766\text{e-}1 & -1.0930\text{e-}1 \end{bmatrix}.$$

The MATLAB code is developed on the basis of Lyapunov iterations which converges to the required solutions of CCSAREs (5.30) with an accuracy of $1.0\text{e} - 12$ order after 76 iterations. It should be noted that the incentive strategy (5.13) that will be announced by the leader can calculated at this time.

Through this incentive, the follower will select the same strategy-set by applying the standard LQ theory,

$$u_{Fij}(t) = -\bar{R}_{Fij}^{-1}(P_{Fi}\bar{B}_{Fij} + \bar{S}_{ij})^T x(t). \tag{5.39}$$

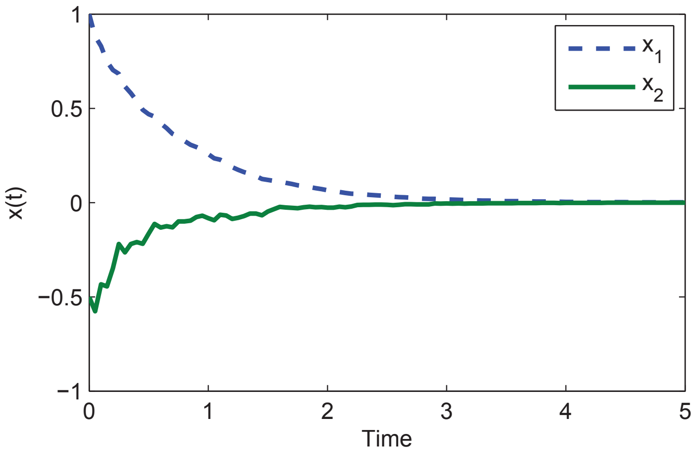


Fig. 5.2: Trajectory of the state $x(t)$.

Namely, it can be confirmed that the followers take the Nash equilibrium solution with H_∞ constraint eventually. Finally, the response is depicted in Fig. 5.2. As a result, it can be observed that the state attains the mean-square stable.

5.6 Conclusion

This chapter investigates the incentive Stackelberg game for continuous-time stochastic systems. The multiple leaders and multiple followers in an incentive Starkberg game are the main features of this chapter. For this game, leaders implement Nash equilibrium in their own group. In the followers' group, players are supposed to be non-cooperative; subsequently, Nash equilibrium is investigated.

A deterministic disturbance input, multiple non-cooperative leaders, and multiple non-cooperative followers have also been considered in this chapter which is different from the previous chapters. An incentive structure is developed in such a way that leaders achieve Nash solution attenuating the disturbance under H_∞ constraint. Simultaneously, followers achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance is considered. As far as we know, this is the first time study for a linear stochastic system with H_∞ constraint involving such multi-leader-follower complicated structure.

This chapter studies the most common linear quadratic (LQ) optimal control in the game problems. Results based on only infinite time domains are shown in this chapter. A computational algorithm based on Lyapunov iterations is used to solve some matrix-valued algebraic equations. In this study, some CCAREs and MAEs in the infinite-horizon case were established, so that the incentive Stackelberg strategy can be achieved with an easy numerical simulation. Several theorems and lemmas are designed to study the incentive Stackelberg game problems. In order to demonstrate the effectiveness of the proposed method, a numerical example is demonstrated with the state trajectory figure.

However, the structure of the game is complex, this chapter only investigates the incentive Stackelberg game with state feedback. This complex structure with output feedback will make the game more interesting in further research. In addition, Markov jump parameters in this study have not been fully investigated and this will be converged in the future study. Finally, a lot of matrix variables are needed to solve the multi-player game problem that takes up a lot of computer memory. Therefore, the number of players can be increased in such a limit so that the computer memory does not fail.

Chapter 6

H_∞-Constrained Pareto Optimal Strategy for Stochastic LPV Systems with Multiple Decision Makers

This chapter is based on a previously published article [Ahmed et al. (2018)].

6.1 Introduction

The dynamic games and the many related applications in practical control problems have been widely investigated by several researchers (see, e.g., [Başar and Bernhard (2008), Başar and Olsder (1999), Engwerda (2005)], and references therein). Starting from the deterministic cases for the continuous and discrete-time, systems have been extended to the stochastic case. Moreover, recent advances in the game theory for a class of stochastic systems revisit the robust and multi-objective control problems [Chen and Zhang (2004), Zhang et al. (2008), Huang et al. (2008)]. The stochastic dynamic games can be solved even if the systems dynamics include a noisy process known as the Wiener process. Additionally, the influence of the deterministic disturbance in the systems model can be reduced by applying the H_∞ control method. Although these results comprise an elegant theory and despite the possibility of obtaining an equilibrium strategy-set, the treatment of uncertainties in the systems state equations continue to remain an issue to be considered in the dynamic games. In other words, the essential core implementation of this strategy-set will determine the notations of the system's unmodeled dynamics.

In robust control design and synthesis, there exists a wide class of dynamic systems that are subject to arbitrary smooth or discontinuous variations in the systems uncertainties. In order to capture these variations in the parameters, linear parameter varying (LPV) systems are reliable to model a large number of parameter variations and these systems offer adequate mathematical models for numerous and phenomena [Apkarian et al. (1995)]. The gain-scheduled design technique is among the most popular methods for designing a robust control by adjusting the scheduling parameters, which describe the changes in plant dynamics. With the maturity of gain-scheduling (GS) control, several results have been reported across various control fields [Sato (2011), Mahmoud (2002), Ku and Wu (2015), Rotondo (2015)]. Stability and H_∞-filtering problems for a class of LPV discrete-time systems in which the state-space matrices depend affinely on time-varying parameters have been investigated in [Mahmoud (2002)]. A new design method for Gain-Scheduled Output Feedback (GSOF) controllers for continuous-time LPV systems via parameter-dependent Lyapunov functions has been tackled in [Sato (2011)]. A GS controller design method has been proposed for LPV stochastic systems subject to H_∞ performance constraint [Ku and Wu (2015)]. Linear quadratic control (LQC) using linear matrix inequalities (LMIs) to LPV systems has been extended [Rotondo (2015)]. Although fruitful results on LPV control design can be found in recent publications, most of them are focused on one control input as a unique decision maker. Considering the fact that the game theory in robust control has become a priority research topic, investigation of the stochastic dynamic games for LPV systems with multiple decision makers is extremely attractive.

Chapter 6 discusses the Pareto optimal strategy for stochastic LPV system with multiple decision makers. In the dynamic game of uncertain stochastic systems, multiple participants can be used for more realistic plants. The system includes disturbances that are attenuated under the H_∞ constraint. This chapter can be seen as an extension of [Mukaidani (2017a)]. This is because the fixed gain controller is also considered here to understand the practical implementation. In this chapter, we design a method for Pareto optimal solution that satisfies the H_∞ norm condition. We redesigned the stochastic bounded real lemma [Ku and Wu (2015)] and the linear quadratic control [Rotondo (2015)] to find the solution. Solvability conditions are established using LMIs. For multiple decision makers, a Pareto optimal strategy-set is designed. The Pareto optimal strategy-set can be found by solving a set of cross-coupling matrix inequalities (CCMI). Academic and practical numerical examples are provided to demonstrate the effectiveness of the proposed model of the LPV system. In the practical point of view, the advantages of the method proposed in this chapter are:

- to alleviate of propagation of uncertainty in stochastic plants;
- to operte linear time-invariant (LTI) plants subject to time-varying parametric uncertainty $\theta(t)$;

- they can be modeled with linear time-varying plants or linearized by nonlinear plants along the trajectory of parameter θ;
- to explain the time-varying nature of the plant through gain-scheduling techniques;
- to process the whole parameter range of the plant with one shot without performing extensive simulation;
- to develop worst-case controller design of linear plants with additive disturbances and plant uncertainties, including problems of disturbance attenuation;
- to minimize multi-objective optimization plant-size and maximize target coverage with Pareto approach.

Notation: The notations used in this Chapter are fairly standard. $\mathbb{E}[\cdot]$ denotes the expectation operator. I_n denotes the $n\times n$ identity matrix. $\|\cdot\|$ denotes the Euclidean norm of a matrix. $L_F^2([0,\ \infty],\ \mathbb{R}^k)$ denotes the space of nonanticipative stochastic process $\phi(t)\in\mathbb{R}^k$ with respect to an increasing σ-algebras $F_t,\ t\geq 0$ satisfying $\mathbb{E}[\int_0^\infty\|\phi(t)\|^2dt]<\infty$.

6.2 Preliminaries

Consider the following stochastic LPV system.

$$dx(t)=[A(\theta(t))x(t)+Bu(t)+Dv(t)]\,dt+A_p(\theta(t))x(t)dw(t),\ x(0)=x^0, \tag{6.1a}$$

$$z(t)=E(\theta(t))x(t), \tag{6.1b}$$

where $x(t)\in\mathbb{R}^n$ denotes the state vector. $u(t)\in\mathbb{R}^m$ denotes the control input. $v(t)\in\mathbb{R}^{n_v}$ denotes the external disturbance. $z(t)\in\mathbb{R}^{n_z}$ denotes the controlled output. $w(t)\in\mathbb{R}$ denotes a one-dimensional standard Wiener process defined in the filtered probability space [Chen and Zhang (2004), Zhang et al. (2008), Huang et al. (2008), Rami and Zhou (2000)]. $\theta(t)\in\mathbb{R}^r$ denotes the time-varying parameters. r is the number of time-varying parameters. It is assumed that the stochastic system (6.1) has a unique strong solution $x(t)=\tilde{x}(t,x(0))$ [Arapostathis et al. (2010)]. The coefficient matrices $A(\theta(t))$ and $A_p(\theta(t))$ are parameter dependent matrices and these matrices can be expressed as

$$\begin{bmatrix}A(\theta(t)) & A_p(\theta(t))\end{bmatrix}=\sum_{k=1}^{M}\alpha_k(t)\begin{bmatrix}A_k & A_{pk}\end{bmatrix}, \tag{6.2a}$$

$$E(\theta(t))=\sum_{k=1}^{M}\alpha_k(t)E_k, \tag{6.2b}$$

where $\alpha_k(t) \geq 0,\ \sum_{k=1}^{M} \alpha_k(t) = 1,\ M = 2^r$.

It should be noted that for simplifying the context of this Chapter, the above-mentioned descriptions are used. Furthermore, many problems involving the synthesis of controllers in the case of a constant B can be reformulated as a varying $B(\theta)$ involving augmented plants [Rotondo (2015)]. The following definition on stochastic stability will be needed.

Definition 6.1. *[Ku and Wu (2015)] A stochastic LPV autonomous system governed by the Itô differential equotion* (6.1) *without $u(t)$ is called mean square stable if the trajectories satisfy*

$$\lim_{t\to\infty} \mathbb{E}[\|x(t)\|^2] = 0,$$

for any initial condition.

The H_∞ norm, which is an essential assumption, is introduced in [Ku and Wu (2015)].

Definition 6.2. *The H_∞ norm of stochastic LPV autonomous system* (6.1) *with mean square stable is given by*

$$\|L\|_\infty^2 = \sup_{\substack{v \in L^2_{\mathscr{F}}([0,\ \infty),\ \mathbb{R}^{n_v}),\\ v \neq 0,\ x^0 = 0}} \frac{J_z}{J_v}, \tag{6.3}$$

where

$$J_z := \mathbb{E}\left[\int_0^\infty \|z(t)\|^2 dt\right],$$
$$J_v := \mathbb{E}\left[\int_0^\infty \|v(t)\|^2 dt\right].$$

Lemma 6.1. *Let us consider an autonomous system such that $u(t) \equiv 0$. For a given attenuation performance level $\gamma > 0$, if there exists matrix $Z > 0$ satisfying the following linear matrix inequalities (LMIs)* (6.4)*, the stochastic LPV system* (6.1) *is mean square stable with $\|L\|_\infty < \gamma$ under $x^0 = 0$.*

$$\Gamma_k := \begin{bmatrix} ZA_k + A_k^T Z & ZD & A_{pk}^T Z & E_k^T \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ ZA_{pk} & 0 & -Z & 0 \\ E_k & 0 & 0 & -I_{n_z} \end{bmatrix} < 0,\ k = 1,\ \cdots, M. \tag{6.4}$$

Moreover, the worst-case disturbance is given by

$$v^*(t) = \gamma^{-2} D^T Z x(t). \tag{6.5}$$

Proof. The following parameter independent Lyapunov function is chosen:

$$V_v(x(t)) = V_v(x) = x^T(t)Zx(t), \tag{6.6}$$

where $Z = Z^T > O$.

The following derivation can be obtained along the trajectories of the stochastic LPV system (6.1) by using Itô's formula.

$$dV_v(x) = LV_v(x)dt + 2x^T A_p^T(\theta)Zxdw(t), \tag{6.7}$$

where

$$LV_v(x) := x^T \left(ZA(\theta) + A^T(\theta)Z + A_p^T(\theta)ZA_p(\theta)x\right) + 2x^T ZDv.$$

In this case, we have the following.

$$\begin{aligned} &LV_v(x) - \gamma^2\|v\|^2 + \|z\|^2 \\ &\quad = x^T\Phi(\theta)x - \gamma^2\left(v - \gamma^{-2}D^T Zx\right)^T\left(v - \gamma^{-2}D^T Zx\right), \end{aligned} \tag{6.8}$$

where

$$\Phi(\theta) := ZA(\theta) + A^T(\theta)Z + A_p^T(\theta)ZA_p(\theta) + E^T(\theta)E(\theta) + \gamma^{-2}ZDD^T Z.$$

Hence, if

$$v(t) = v^*(t) = -\gamma^{-2}D^T Zx(t)$$

holds, we have

$$LV_v(x) - \gamma^2\|v\|^2 + \|z\|^2 \leq x^T\Phi(\theta)x. \tag{6.9}$$

Thus, the worst-case disturbance (6.5) can be obtained. Moreover, $\Phi(\theta) < 0$ is equivalent to the following LMI by using the Schur complement.

$$\Lambda(\theta(t)) := \begin{bmatrix} \Psi(\theta) & ZD \\ D^T Z & -\gamma^2 I_{n_v} \end{bmatrix}, \tag{6.10}$$

where

$$\Psi(\theta) := ZA(\theta) + A^T(\theta)Z + A_p^T(\theta)ZA_p(\theta) + E^T(\theta)E(\theta).$$

By integrating and taking the expectation both sides of the equality (6.7) from 0 to t_f, the following equation holds under the assumption that $LV_v(x) - \gamma^2\|v\|^2 + \|z\|^2 < x^T\Phi(\theta)x < 0$ from (6.8):

$$\mathbb{E}\left[\int_0^{t_f} dV_v(x)\right] = V_v(x(t_f)) - V_v(x(0)) = V_v(x(t_f))$$

$$< \mathbb{E}\left[\int_0^{t_f} \gamma^2\|v(t)\|^2 + \|z(t)\|^2\right]. \tag{6.11}$$

From (6.11), it is easy to see that if $\Lambda(\theta) < 0$ then $\|L\|_\infty < \gamma$ because $J_z < \gamma^2 J_v$ and $V_v(x(t_f)) > 0$ as $t_f \to 0$. Therefore, $\Lambda(\theta) < 0$ is considered. By applying the Schur complement, inequality (6.10) is seen to be equivalent to the LMI as

$$\Lambda(\theta(t)) := \begin{bmatrix} \Psi(\theta) & ZD \\ D^T Z & -\gamma^2 I_{n_v} \end{bmatrix} < 0,$$

which implies

$$\begin{bmatrix} ZA(\theta)+A^T(\theta)Z+A_p^T(\theta)ZA_p(\theta)+E^T(\theta)E(\theta) & ZD \\ D^T Z & -\gamma^2 I_{n_v} \end{bmatrix} < 0,$$

$$\begin{bmatrix} ZA(\theta)+A^T(\theta)Z & ZD \\ D^T Z & -\gamma^2 I_{n_v} \end{bmatrix} + \begin{bmatrix} A_p^T(\theta)ZA_p(\theta)+E^T(\theta)E(\theta) & 0 \\ 0 & 0 \end{bmatrix} < 0,$$

$$\begin{bmatrix} ZA(\theta)+A^T(\theta)Z & ZD \\ D^T Z & -\gamma^2 I_{n_v} \end{bmatrix} - \begin{bmatrix} A_p^T(\theta)Z & E^T(\theta) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -Z & 0 \\ 0 & -I_{n_z} \end{bmatrix}^{-1} \begin{bmatrix} ZA_p(\theta) & 0 \\ E(\theta) & 0 \end{bmatrix} < 0,$$

$$\left[\begin{array}{cc|cc} ZA(\theta)+A^T(\theta)Z & ZD & A_p^T(\theta)Z & E^T(\theta) \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ \hline ZA_p(\theta) & 0 & -Z & 0 \\ E(\theta) & 0 & 0 & -I_{n_z} \end{array}\right] < 0,$$

$$\begin{bmatrix} ZA(\theta)+A^T(\theta)Z & ZD & A_p^T(\theta)Z & E^T(\theta) \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ ZA_p(\theta) & 0 & -Z & 0 \\ E(\theta) & 0 & 0 & -I_{n_z} \end{bmatrix} < 0. \tag{6.12}$$

If the parameter dependent coefficient matrices are changed by applying (6.2), the inequality (6.12) can be written as follows:

$$\Gamma_k := \begin{bmatrix} ZA_k+A_k^T Z & ZD & A_{pk}^T Z & E_k^T \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ ZA_{pk} & 0 & -Z & 0 \\ E_k & 0 & 0 & -I_{n_z} \end{bmatrix} < 0,\ k = 1, \cdots, M.$$

Thus, the proof of Lemma 6.1 is completed. □

On the other hand, the standard linear quadratic control (LQC) problem for the stochastic LPV system with $v(t) \equiv 0$ or $D(\theta(t)) \equiv 0$ is given [Rotondo (2015)].

Definition 6.3. *Let us consider the stochastic LPV system with $v(t) \equiv 0$ in (6.1). The following cost performance is defined by*

$$J(u,\, x^0) = \mathbb{E}\left[\int_0^{\infty} [x^T(t)Qx(t) + u^T(t)Ru(t)]dt\right], \tag{6.13}$$

where $Q = Q^T > 0$ and $R = R^T > 0$.

In this situation, the LQC problem is to find a state feedback control

$$u(t) = K(\theta(t))x(t) = \sum_{k=1}^{M} \alpha_k(t) K_k x(t) \tag{6.14}$$

such that the quadratic cost functional (6.13) *is minimized.*

Lemma 6.2. *If there exists the matrix $X > 0$ and Y_k, $k = 1, \cdots, M$ satisfying the LMI* (6.15)*:*

$$\begin{bmatrix} \Xi_k & X & Y_k^T & XA_{pk}^T \\ X & -Q^{-1} & 0 & 0 \\ Y_k & 0 & -R^{-1} & 0 \\ A_{pk}X & 0 & 0 & -X \end{bmatrix} < 0, \tag{6.15a}$$

$$\begin{bmatrix} \Xi_{kl} & X & Y_{kl}^T & XA_{pkl}^T \\ X & -\frac{1}{2}Q^{-1} & 0 & 0 \\ Y_{kl} & 0 & -\frac{1}{2}R^{-1} & 0 \\ A_{pkl}X & 0 & 0 & -\frac{1}{2}X \end{bmatrix} < 0, \tag{6.15b}$$

where

$$\begin{aligned} & k < l,\ k = 1,\ \cdots, M, \\ & \Xi_k = A_k X + XA_k^T + BY_k + Y_k^T B^T, \\ & \Xi_{kl} = (A_k + A_l)X + X(A_k + A_l)^T + BY_k + Y_k^T B^T + BY_l + Y_l^T B^T, \\ & K_k = Y_k X^{-1},\ K_l = Y_l X^{-1}, \\ & A_{pkl} := \frac{1}{2}(A_{pk} + A_{pl}), \\ & Y_{kl} := \frac{1}{2}(Y_k + y_l), \end{aligned}$$

then

$$u(t) = \sum_{k=1}^{M} \alpha_k(t) K_k x(t) = \sum_{k=1}^{M} \alpha_k(t) Y_k X^{-1} x(t), \tag{6.16a}$$

$$J(u,\ x^0) < \mathbb{E}\left[x^T(0) X^{-1} x(0)\right]. \tag{6.16b}$$

Proof. First, the following parameter independent Lyapunov function is introduced:

$$V_u(x(t)) = V_u(x) = x^T(t) P x(t), \tag{6.17}$$

where $P = P^T > 0$ with $P = X^{-1}$.

Let us consider the closed-loop stochastic LPV system With the control (6.14). By using a similar technique to that in [Mukaidani (2009)], if there exists $P > 0$ such that

$$P(A(\theta)+BK(\theta))+(A(\theta)+BK(\theta))^T P \\ +A_p^T(\theta)PA_p(\theta)+Q+K^T(\theta)RK(\theta) < 0, \tag{6.18}$$

then, the following equation holds.

$$J(u,\, x^0) < \mathbb{E}\left[x^T(0)Px(0)\right] \tag{6.19}$$

In this case, by rearranging equation (6.18), we have the following.

$$\sum_{k=1}^{M} \alpha_k^2 \left(P(A_k+BK_k)+(A_k+BK_k)^T P + A_{pk}^T PA_{pk}+Q+K_k^T RK_k\right) \\ +\sum_{k=1}^{M-1}\sum_{l=k+1}^{M} \alpha_k\alpha_l \left(PG_{kl}+G_{kl}^T P+H_{kl}+2Q+T_{kl}\right) < 0, \tag{6.20}$$

where

$$G_{kl} := A_k+BK_k+A_l+BK_l,$$
$$H_{kl} := A_{pk}^T PA_{pl}+A_{pl}^T PA_{pk},$$
$$T_{kl} := K_k^T RK_l+K_l^T RK_k.$$

On the other hand, applying Schur complement on the inequality (6.15a) and (6.15b), the following matrix inequalities hold.

$$\left[\begin{array}{cc|cc} A_kX+XA_k^T+BY_k+Y_k^T B^T & X & Y_k^T & XA_{pk}^T \\ X & -Q^{-1} & 0 & 0 \\ \hline Y_k & 0 & -R^{-1} & 0 \\ A_{pk}X & 0 & 0 & -X \end{array}\right] < 0, \tag{6.21a}$$

$$\left[\begin{array}{cc|cc} (A_k+A_l)X+X(A_k+A_l)^T+BY_k & & & \\ +Y_k^T B^T+BY_l+Y_l^T B^T & X & Y_{kl}^T & XA_{pkl}^T \\ X & -\frac{1}{2}Q^{-1} & 0 & 0 \\ \hline Y_{kl} & 0 & -\frac{1}{2}R^{-1} & 0 \\ A_{pkl}X & 0 & 0 & -\frac{1}{2}X \end{array}\right] < 0, \tag{6.21b}$$

where

$$K_k := Y_kX^{-1},$$
$$K_l := Y_lX^{-1},$$

$$A_{pkl} := \frac{1}{2}(A_{pk} + A_{pl}),$$
$$Y_{kl} := \frac{1}{2}(Y_k + y_l).$$

Or,

$$\begin{bmatrix} A_kX + XA_k^T \\ +BY_k + Y_k^TB^T & X \\ X & -Q^{-1} \end{bmatrix} - \begin{bmatrix} Y_k^T & XA_{pk}^T \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -R^{-1} & 0 \\ 0 & -X \end{bmatrix}^{-1} \begin{bmatrix} Y_k & 0 \\ A_{pk}X & 0 \end{bmatrix} < 0, \tag{6.22a}$$

$$\begin{bmatrix} (A_k + A_l)X + X(A_k + A_l)^T + BY_k + Y_k^TB^T + BY_l + Y_l^TB^T & X \\ X & -\frac{1}{2}Q^{-1} \end{bmatrix}$$
$$- \begin{bmatrix} Y_{kl}^T & XA_{pkl}^T \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}R^{-1} & 0 \\ 0 & -\frac{1}{2}X \end{bmatrix}^{-1} \begin{bmatrix} Y_{kl} & 0 \\ A_{pkl}X & 0 \end{bmatrix} < 0, \tag{6.22b}$$

where

$$K_k = Y_kX^{-1},\ K_l = Y_lX^{-1},\ A_{pkl} := \frac{1}{2}(A_{pk} + A_{pl}),\ Y_{kl} := \frac{1}{2}(Y_k + y_l).$$

Or,

$$A_kX + XA_k^T + BY_k + Y_k^TB^T + XQ^{-1}X + Y_k^TRY_k + A_{pk}^TXA_{pk} < 0, \tag{6.23a}$$
$$(A_k + A_l)X + X(A_k + A_l)^T + BY_k + Y_k^TB^T + BY_l + Y_l^TB^T$$
$$+ X\frac{1}{2}Q^{-1}X + \frac{1}{2}Y_{kl}^TRY_{kl} + \frac{1}{2}A_{pkl}^TXA_{pkl} < 0, \tag{6.23b}$$

where

$$K_k = Y_kX^{-1},\ K_l = Y_lX^{-1},\ A_{pkl} := \frac{1}{2}(A_{pk} + A_{pl}),\ Y_{kl} := \frac{1}{2}(Y_k + y_l).$$

Pre- and post- multiplying both sides on inequality (6.23) by P yields,

$$P(A_k + BK_k) + (A_k + BK_k)^TP + A_{pk}^TPA_{pk} + Q + K_k^TRK_k < 0, \tag{6.24a}$$
$$PG_{kl} + G_{kl}^TP + 2A_{pkl}^TPA_{pkl} + 2Q + 2\left(\frac{K_k + K_l}{2}\right)^T \cdot R \cdot \frac{K_k + K_l}{2} < 0, \tag{6.24b}$$

where $K_k = Y_kX^{-l} = Y_kP$.

Furthermore, it is well known that the following inequalities hold.

$$2A_{pkl}^TPA_{pkl} \geq H_{kl}, \tag{6.25a}$$
$$2\left(\frac{K_k + K_l}{2}\right)^T \cdot R \cdot \frac{K_k + K_l}{2} \geq T_{kl}. \tag{6.25b}$$

Hence, inequality (21b) can be changed as follows.

$$PG_{kl} + G_{kl}^TP + H_{kl} + 2Q + T_{kl} < 0. \tag{6.26}$$

Thus, if inequalities (6.24a) and (6.26) are satisfied, then inequality (6.20) holds. In other words, if inequalities (6.15a) and (6.15b) are satisfied, then (6.18) holds and this inequality implies the cost bound (6.16b). □

It should be noted that the obtained result corresponding to Lemma 6.1 is not a necessary and sufficient condition as compared with the existing result of [Apkarian et al. (1995)] but the conditions of (6.15) are the sufficient conditions.

6.3 Problem Formulation

Consider a stochastic LPV system governed by Itô differential equation with multiple decision makers defined by

$$dx(t) = \left[A(\theta(t))x(t) + \sum_{j=1}^{N} B_j u_j(t) + Dv(t)\right] dt + A_p(\theta(t))x(t)dw(t),\ x(0) = x^0, \tag{6.27a}$$

$$z(t) = \begin{bmatrix} E(\theta(t))x(t) \\ G_1 u_1(t)) \\ \vdots \\ G_N u_N(t) \end{bmatrix}, \tag{6.27b}$$

where $u_i(t) \in \mathbb{R}^{m_i},\ i = 1,\ \cdots,N$ denote the i-th control inputs. Other variables are defined by stochastic equation (6.1). It should be noted that G_i does not depend on the time-varying parameter because the controlled output can be chosen by the controller designer. Hence, without loss of generality, it may be assumed that G_i is a constant matrix.

Assumption 6.1. $G_i^T G_i = I_m,\ i = 1,\ \cdots,N,\ G_i \in \mathbb{R}^{g_i \times m_i}$.

The cost performances are defined by

$$J_v(u_1,\ \cdots,\ u_N,\ v,\ x^0) = \mathbb{E}\left[\int_0^\infty \left[\gamma^2\|v(t)\|^2 - \|z(t)\|^2\right] dt\right], \tag{6.28a}$$

$$J_i(u_1,\ \cdots,\ u_N,\ v,\ x^0) = \mathbb{E}\left[\int_0^\infty \left[x^T(t)Q_i x(t) + u_i^T(t)R_i u_i(t)\right] dt\right], \tag{6.28b}$$

where $Q_i = Q_i^T > 0$ *and* $R_i = R_i^T > 0$.

The infinite horizon gain-scheduled H_∞ constraint Pareto optimal control strategy [Engwerda (2005)] for the stochastic LPV system (6.27) is described as follows.

Definition 6.4. *For given* $\gamma > 0,\ v(t) \in L_F^2([0,\ \infty),\ \mathbb{R}^{m_v})$*, find a state feedback strategy-set* $u_i(t) = u_i^*(t) \in L_F^2([0,\ \infty)),\ i = 1,\ \cdots,N$ *such that*

(i) The trajectory of the closed-loop system of stochastic system (6.27) *satisfies*

$$0 \le J_v(u_1^*,\ \cdots,u_N^*,\ v^*,\ x^0) \le J_v(u_1^*,\ \cdots,u_N^*,\ v,\ x^0), \tag{6.29}$$

where $v^*(t)$ is the worst-case disturbance.
(ii) When the worst-case disturbance $v^*(t)$ is implemented in (6.27)*, $u_i(t),\ i = 1,\ \cdots,N$ minimizes a sum of the cost of function of all decision makers denoted by*

$$J(u_1,\ \cdots,u_N,\ v^*,\ x^0) = \sum_{j=1}^{N} r_j J_j(u_1,\ \cdots,u_N,\ v^*,\ x^0), \tag{6.30}$$

where $0 < r_i < 1,\ \sum_{j=1}^{N} r_j = 1$ for some r_i.

It should be noted that Pareto optimality is not necessarily equivalent to the weighted sum minimization [Engwerda (2005)].

For the next we will establish the solution of the above-mentioned problem which is called the H_∞ constraint Pareto optimal strategy.

6.4 Gain-Scheduled H_∞ Constrained Pareto Optimal Solution

6.4.1 Main Result

We now in give the main contribution of this Chapter.

Theorem 6.1. *Let us consider the stochastic LPV system* (6.27) *with multiple decision makers $u_i(t)$ and the disturbance $v(t)$. For a given attenuation performance level $\gamma > 0$, assume that there exists a solution set for the real symmetric matrices $X > 0$, Y_k, Y_l and $Z > 0$ such that the following CCMIs are satisfied:*

$$\begin{bmatrix} \Xi_k & X & Y_k^T & XA_{pk}^T \\ X & -\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_k & 0 & -\boldsymbol{R}^{-1} & 0 \\ A_{pk}X & 0 & 0 & -X \end{bmatrix} < 0, \tag{6.31a}$$

$$\begin{bmatrix} \Xi_{kl} & X & Y_{kl}^T & XA_{pkl}^T \\ X & -\frac{1}{2}\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_{kl} & 0 & -\frac{1}{2}\boldsymbol{R}^{-1} & 0 \\ A_{pkl}X & 0 & 0 & -\frac{1}{2}X \end{bmatrix} < 0, \tag{6.31b}$$

$$\begin{bmatrix} Z\boldsymbol{A}_{-Fk} + \boldsymbol{A}_{-Fk}^T Z & ZD & A_{pk}^T Z & \boldsymbol{E}_k^T \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ ZA_{pk} & 0 & -Z & 0 \\ \boldsymbol{E}_k & 0 & 0 & -I_{n_{zg}} \end{bmatrix} < 0, \tag{6.31c}$$

where

$$k < l,\ k = 1,\ \cdots,M,$$

$$\boldsymbol{\Xi}_k := \boldsymbol{A}_k X + X\boldsymbol{A}_k^T + BY_k + Y_k^T B^T,$$
$$\boldsymbol{\Xi}_{kl} := (\boldsymbol{A}_k + \boldsymbol{A}_l)X + X(\boldsymbol{A}_k + \boldsymbol{A}_l)^T + BY_k + Y_k^T B^T + BY_l + Y_l^T B^T,$$
$$A_{-Fk} := A_k + \sum_{j=1}^{N} B_j K_{jk},$$
$$\boldsymbol{A}_k := A_k + DF,$$
$$E_k := \begin{bmatrix} E_k^T & (G_1 K_{1k})^T & \cdots & (G_N K_{Nk})^T \end{bmatrix}^T,$$
$$n_{zg} := n_z + \sum_{j=1}^{N} g_i,$$
$$K_k := Y_k X^{-1},$$
$$K_l := Y_l X^{-1},$$
$$Y_{kl} := \frac{1}{2}(Y_k + Y_l),$$
$$F := \gamma_{-2} D^T Z,$$
$$\boldsymbol{Q} := \sum_{j=1}^{N} r_j Q_j,$$
$$\boldsymbol{R} := \boldsymbol{block\ diag}(r_1 R_1 \ \cdots \ r_N R_N).$$

Then, the following controllers comprise the Pareto optimal strategy-set.

$$u(t) = K_i(\theta)x(t) = \sum_{k=1}^{M} \alpha_k(t) K_{ik} x(t) = \sum_{k=1}^{M} \alpha_k(t) Y_{ki} X^{-1} x(t), \tag{6.32}$$

where $K_k := \begin{bmatrix} K_{1k}^T & \cdots & K_{Nk}^T \end{bmatrix}^T$. *Furthermore, the optimal cost bound is given by*

$$J(u_1^*, \cdots, u_N^*, v^*, x^0) \le \mathbb{E}\left[x^T(0) X^{-1} x(0)\right], \tag{6.33}$$

where $v^*(t) = Fx(t)$.

Proof. First, the H_∞ constraint condition is investigated. The Pareto optimal strategy-set (6.32) is applied to original stochastic LPV system (6.27) and we have the following closed loop stochastic system.

$$dx(t) = \left[\left(A(\theta(t)) + \sum_{j=1}^{N} B_j K_j(\theta(t))\right) x(t) + Dv(t)\right] dt + A_p(\theta(t)) x(t) dw(t),\ x(0) = x^0, \tag{6.34a}$$

$$z(t) = \boldsymbol{E}(\theta(t)) x(t). \tag{6.34b}$$

Hence, by the term-wise comparison between (6.1) and (6.34), we have

$$A(\theta) \leftarrow A(\theta) + \sum_{j=1}^{N} B_j K_j(\theta) = \sum_{k=1}^{M} \alpha_k \boldsymbol{A}_{-Fk}, \tag{6.35a}$$

$$E(\theta) \leftarrow \boldsymbol{E}(\theta) = \sum_{k=1}^{M} \alpha_k \boldsymbol{E}_k. \tag{6.35b}$$

Thus, by applying Lemma 6.1 to this problem, LMI (6.31c) can be obtained. Moreover, the condition of the existence of the Pareto optimal strategy-set is derived. The following LQC problem is considered.

$$\min J(u,\ v,\ x^0) = \min_u \mathbb{E}\left[\int_0^{\infty} \left[x^T(t)\boldsymbol{Q}x(t) + u^T(t)\boldsymbol{R}u(t)\right] dt\right], \tag{6.36}$$

where $u(t) = \begin{bmatrix} u_1^T & \cdots & u_N^T \end{bmatrix}^T$ such that

$$dx(t) = [(A(\theta(t)) + DF)x(t) + Bu(t)]\,dt + A_p(\theta(t))x(t)dw(t),\ x(0) = x^0, \tag{6.37}$$

where $B = \begin{bmatrix} B_1 & \cdots & B_N \end{bmatrix}$.

Hence, as the similar step of the H_∞ constraint problem, by the term-wise comparison between (6.1) and (6.36) with (6.52), we have

$$A(\theta) \leftarrow A(\theta(t)) + DF + BK(\theta(t)) = \sum_{k=1}^{M} \alpha_k \boldsymbol{A}_k, \tag{6.38a}$$

$$Q \leftarrow \boldsymbol{Q},\ R \leftarrow \boldsymbol{R}. \tag{6.38b}$$

Therefore, by applying Lemma 6.2 to this problem, CCMIs (6.31a) and (6.31b) can be obtained. □

It should be noted that the Existence of the solutions in inequality (6.31) is not guaranteed and these conditions are conservative in general. It may also be pointed out that a weakly sufficient condition (dense in the set of all Pareto equilibria) that usually asserts the statement based on the Arrow-Barankin-Blackwell theorem [Arrow et al. (1953)].

6.4.2 Numerical Algorithm for Solving CCMIs

In order to construct the Pareto optimal strategy-set of (6.32), we must solve the CCMIs (6.31). It should be noted that since these matrix inequalities are coupled, it is very complicated if an ordinary scheme such as Newton's method is applied. In this section, a numerical algorithm via the semidefinite programming problem (SDP) is considered.

Step 1. As the first step, any weight r_i for the cost function (6.30) and solve the following SDP.

$$\text{minimize }\ \mathbf{Tr}\left[\alpha^{(0)}\right], \tag{6.39}$$

subject to

$$\begin{bmatrix} \boldsymbol{\Xi}_k^{(0)} & X^{(0)} & Y_k^{(0)T} & X^{(0)}A_{pk}^T \\ X^{(0)} & -\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_k^{(0)} & 0 & -\boldsymbol{R}^{-1} & 0 \\ A_{pk}X^{(0)} & 0 & 0 & -X^{(0)} \end{bmatrix} < 0, \tag{6.40a}$$

$$\begin{bmatrix} \boldsymbol{\Xi}_{kl}^{(0)} & X^{(0)} & Y_{kl}^{(0)T} & X^{(0)}A_{pkl}^T \\ X^{(0)} & -\frac{1}{2}\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_{kl}^{(0)} & 0 & -\frac{1}{2}\boldsymbol{R}^{-1} & 0 \\ A_{pkl}X^{(0)} & 0 & 0 & -\frac{1}{2}X^{(0)} \end{bmatrix} < 0, \tag{6.40b}$$

$$\begin{bmatrix} -\alpha^{(0)} & x^T(0) \\ x(0) & -X^{(0)} \end{bmatrix} < 0, \tag{6.40c}$$

where

$$\begin{aligned}
&k < l,\ k = 1,\ \cdots,M, \\
&\boldsymbol{\Xi}_k^{(0)} := A_k X^{(0)} + X^{(0)} A_k^T + BY_k^{(0)} + Y_k^{(0)T} B^T, \\
&\boldsymbol{\Xi}_{kl}^{(0)} := (A_k + A_l)X^{(0)} + X^{(0)}(A_k + A_l)^T + BY_k^{(0)} + Y_k^{(0)T} B^T + BY_l^{(0)} + Y_l^{(0)T} B^T, \\
&Y_{kl}^{(0)} := \frac{1}{2}(Y_k^{(0)} + Y_l^{(0)}), \\
&K_k^{(0)} := Y_k^{(0)}[X^{(0)}]^{-1}, \\
&K_l^{(0)} := Y_l^{(0)}[X^{(0)}]^{-1}.
\end{aligned}$$

Choose any γ and solve $Z^{(0)}$, where

$$\begin{aligned}
&F^{(0)} := \gamma^{-2} D^T Z^{(0)}, \\
&Z^{(0)}\bar{A} + \bar{A}^T Z^{(0)} + \bar{A}_p^T Z^{(0)} \bar{A}_p + \gamma^{-2} Z^{(0)} D D^T Z^{(0)} + \bar{E}^T \bar{E} = 0, \\
&\bar{A} := \frac{1}{N}\sum_{k=1}^{M} A_k, \\
&\bar{A}_p := \frac{1}{N}\sum_{k=1}^{M} A_{pk}, \\
&\bar{E} := \frac{1}{N}\sum_{k=1}^{M} E_k.
\end{aligned}$$

Step 2. Solve the following SDP.

$$\text{minimize}\ \ \mathbf{Tr}\left[\alpha^{(p)}\right], \tag{6.41}$$

subject to

$$\begin{bmatrix} \boldsymbol{\Xi}_k^{(p)} & X^{(p)} & Y_k^{(p)T} & X^{(p)}A_{pk}^T \\ X^{(p)} & -\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_k^{(p)} & 0 & -\boldsymbol{R}^{-1} & 0 \\ A_{pk}X^{(p)} & 0 & 0 & -X^{(p)} \end{bmatrix} < 0, \tag{6.42a}$$

$$\begin{bmatrix} \boldsymbol{\Xi}_{kl}^{(p)} & X^{(p)} & Y_{kl}^{(p)T} & X^{(p)}A_{pkl}^T \\ X^{(p)} & -\frac{1}{2}\boldsymbol{Q}^{-1} & 0 & 0 \\ Y_{kl}^{(p)} & 0 & -\frac{1}{2}\boldsymbol{R}^{-1} & 0 \\ A_{pkl}X^{(p)} & 0 & 0 & -\frac{1}{2}X^{(p)} \end{bmatrix} < 0, \tag{6.42b}$$

$$\begin{bmatrix} -\alpha^{(p)} & x^T(0 \\ x(0) & -X^{(p)} \end{bmatrix} < 0, \tag{6.42c}$$

where

$$\begin{aligned}
&p = 1,\ 2,\ \cdots;\ k < l,\ k = 1,\ \cdots,M,\\
&\boldsymbol{\Xi}_k^{(p)} := \boldsymbol{A}_k X^{(p)} + X^{(p)}\boldsymbol{A}_k^T + BY_k^{(p)} + Y_k^{(p)T}B^T,\\
&\boldsymbol{\Xi}_{kl}^{(p)} := (\boldsymbol{A}_k + \boldsymbol{A}_l)X^{(p)} + X^{(p)}(\boldsymbol{A}_k + \boldsymbol{A}_l)^T + BY_k^{(p)} + Y_k^{(p)T}B^T + BY_l^{(p)} + Y_l^{(p)T}B^T,\\
&\boldsymbol{A}_k^{(p)} := A_k + DF^{(p-1)} + BK_k^{(p-1)},\\
&Y_{kl}^{(p)} := \frac{1}{2}(Y_k^{(p)} + Y_l^{(p)}),\\
&K_k^{(p-1)} := Y_k^{(p-1)}[X^{(p-1)}]^{-1},\\
&K_l^{(p-1)} := Y_l^{(p-1)}[X^{(p-1)}]^{-1}.
\end{aligned}$$

Step 3. Solve the following SDP.

$$\text{minimize } \mathbf{Tr}\left[x^T(0)Z^{(p)}x(0)\right], \tag{6.43}$$

subject to

$$\begin{bmatrix} \Gamma^{(p)} & Z^{(p)}D & A_{pk}^T Z^{(p)} & E_k^{(p)T} \\ D^T Z^{(p)} & -\gamma^2 I_{n_v} & 0 & 0 \\ Z^{(p)}A_{pk} & 0 & -Z^{(p)} & 0 \\ \boldsymbol{E}_k^{(p)} & 0 & 0 & -I_{n_{zg}} \end{bmatrix} < 0, \tag{6.44}$$

where

$$\begin{aligned}
&p = 1,\ 2,\ \cdots;\ k = 1,\ \cdots,M,\\
&\Gamma^{(p)} := Z^{(p)}\boldsymbol{A}_{-Fk}^{(p)} + \boldsymbol{A}_{-Fk}^{(p)}Z^{(p)},\\
&\boldsymbol{A}_{-Fk}^{(p)} := A_k + BK_k^{(p-1)},\\
&\boldsymbol{E}_k^{(p)} := \left[E_k^T \quad (G_1 K_{1k}^{(p-1)})^T \quad \cdots \quad (G_N K_{Nk}^{(p-1)})^T\right]^T,
\end{aligned}$$

$$F^{(p)} := \gamma^{-2} D^T Z^{(p)}.$$

Step 4. If the algorithm converges, then $X^{(p)} \to X$, $Y_k^{(p)} \to Y_k$ and $Z^{(p)} \to Z$ as $p \to \infty$. They are the solution of CCMIs (6.31), STOP. That is, stop if any norm of the error of difference between the iterative solutions of (6.42), (6.44) and the exact solutions of (6.31) is less than a precision. Otherwise, increment $p \to p+1$ and go to **Step 2**. If the algorithm does not converge, declare the algorithm failed.

It should be noted that convergence of the above algorithm cannot be guaranteed. However, we found that the proposed algorithm works well in practice.

6.5 Fixed Gain Pareto Strategy

In this section, we discuss the fixed gain Pareto optimal strategy-set to enable easy controller design. Consider a stochastic LPV system by Itô differential equation with multiple decision makers defined by

$$dx(t) = \left[A(\theta(t))x(t) + \sum_{j=1}^{N} B_j(\theta(t))u_j(t) + Dv(t) \right] dt + A_p(\theta(t))x(t)dw(t),\ x(0) = x^0, \tag{6.45a}$$

$$u_i(t) = \bar{K}_i x(t). \tag{6.45b}$$

where

$$B_i(\theta(t)) = \sum_{k=1}^{M} \alpha_k(t) B_{ik}.$$

Other variables are defined by stochastic equation (6.27). It should be noted that $u_i(t)$ does not depend on $\theta(t)$ in this section.

Theorem 6.2. *Let us consider stochastic LPV system* (6.45). *For a given attenuation performance level $-\gamma > 0$, assume that there exists a solution set for the real symmetric matrices $\bar{X} > 0$, $\bar{Y}$ and $\bar{Z} > 0$ such that the following CCMIs* (6.46) *are satisfied:*

$$\begin{bmatrix} \bar{\boldsymbol{X}}_{ik} & \bar{X} & \bar{Y}^T & \bar{X}A_{pk}^T \\ \bar{X} & -\boldsymbol{Q}^{-1} & 0 & 0 \\ \bar{Y} & 0 & -\boldsymbol{R}^{-1} & 0 \\ A_{pk}\bar{X} & 0 & 0 & -\bar{X} \end{bmatrix} < 0, \tag{6.46a}$$

$$\begin{bmatrix} \bar{\boldsymbol{X}}_{ikl} & \bar{X} & \bar{Y}^T & \bar{X}A_{pkl}^T \\ \bar{X} & -\frac{1}{2}\boldsymbol{Q}^{-1} & 0 & 0 \\ \bar{Y} & 0 & -\frac{1}{2}\boldsymbol{R}^{-1} & 0 \\ A_{pkl}\bar{X} & 0 & 0 & -\frac{1}{2}\bar{X} \end{bmatrix} < 0, \tag{6.46b}$$

$$\begin{bmatrix} \bar{Z}\boldsymbol{A}_{-Fk}+\boldsymbol{A}_{-Fk}^T\bar{Z} & \bar{Z}D & A_{pk}^T\bar{Z} & \bar{\boldsymbol{E}}_k^T \\ D^T Z & -\gamma^2 I_{n_v} & 0 & 0 \\ \bar{Z}A_{pk} & 0 & -\bar{Z} & 0 \\ \bar{\boldsymbol{E}}_k & 0 & 0 & -I_{n_{zg}} \end{bmatrix} < 0, \tag{6.46c}$$

where

$$\begin{aligned}
&k < l,\ k = 1,\ \cdots,M, \\
&\bar{\boldsymbol{X}}_{ik} := \bar{\boldsymbol{A}}_k\bar{X} + \bar{X}\bar{\boldsymbol{A}}_k^T + B_k\bar{Y} + \bar{Y}^T B_k^T, \\
&\bar{\boldsymbol{X}}_{ikl} := (\bar{\boldsymbol{A}}_k + \bar{\boldsymbol{A}}_l)X + X(\bar{\boldsymbol{A}}_k + \bar{\boldsymbol{A}}_l)^T + B_k\bar{Y} + \bar{Y}^T B_k^T + B_l\bar{Y} + \bar{Y}^T B_l^T, \\
&\bar{\boldsymbol{A}}_{-\bar{F}k} := A_k + \sum_{j=1}^{N} B_{jk}\bar{K}_{jk}, \\
&\bar{\boldsymbol{A}}_k := A_k + D\bar{F}, \\
&B_k := \begin{bmatrix} B_{1k} & \cdots & B_{Nk}, \end{bmatrix} \\
&\bar{\boldsymbol{E}}_k := \begin{bmatrix} E_k^T & (G_1K_{1k})^T & \cdots & (G_NK_{Nk})^T \end{bmatrix}^T, \\
&\bar{F} := \gamma^{-2}D^T\bar{Z}.
\end{aligned}$$

Then, the following controllers comprise the Pareto optimal strategy-set.

$$u^*(t) = \bar{K}x(t) = \bar{Y}\bar{X}^{-1}x(t) = \begin{bmatrix} \bar{K}_1^* \\ \vdots \\ \bar{K}_N^* \end{bmatrix} x(t) = \begin{bmatrix} u_1^*(t) \\ \vdots \\ u_N^*(t) \end{bmatrix}, \tag{6.47}$$

where $u_i^*(t) = \bar{K}_i^* x(t)$.

Furthermore, the optimal cost bound is given by

$$J(u_1^*,\ \cdots,u_N^*,\ v^*,\ x^0) \le \mathbb{E}\left[x^T(0)\bar{X}^{-1}x(0)\right], \tag{6.48}$$

where $v^*(t) = \bar{F}x(t)$.

Proof. By using the similar technique in the previous section, the proof can be completed. Applying (6.45b) to the stochastic LPV system, we have

$$dx(t) = \left[\bar{A}_{-\bar{F}}(\theta(t))x(t) + Dv(t)\right]dt + A_p(\theta(t))x(t)dw(t),\ x(0) = x^0, \tag{6.49a}$$

$$z(t) = \bar{\boldsymbol{E}}(\theta(t))x(t). \tag{6.49b}$$

where

$$\bar{A}_{-\bar{F}}(\theta(t)) := A(\theta(t)) + \sum_{j=1}^{N} B_j(\theta(t))\bar{K}_j = \sum_{k=1}^{M} \alpha_k(t)\bar{\boldsymbol{A}}_{-\bar{F}k},$$

$$\bar{\boldsymbol{E}}(\theta(t)) := \sum_{k=1}^{M} \alpha_k(t)\bar{\boldsymbol{E}}_k.$$

Hence, by the term-wise comparison between (6.1) and (6.49), we have

$$A(\theta) \leftarrow \bar{\boldsymbol{A}}_{-\bar{F}}(\theta) = \sum_{k=1}^{M} \alpha_k \bar{\boldsymbol{A}}_{-\bar{F}k}, \tag{6.50a}$$

$$E(\theta) \leftarrow \bar{\boldsymbol{E}}(\theta) = \sum_{k=1}^{M} \alpha_k \bar{\boldsymbol{E}}_k, \tag{6.50b}$$

Thus, LMI (6.46c) can be obtained. Second, the following LQC problem is considered.

$$\min_u J(u,\, \bar{v}^*,\, x^0) := \min_u \mathbb{E}\left[\int_0^\infty \left[x^T(t)\boldsymbol{Q}x(t) + u^T(t)\boldsymbol{R}u(t)\right] dt\right], \tag{6.51}$$

such that

$$dx(t) = \left[\bar{\boldsymbol{A}}(\theta(t))x(t) + B(\theta(t))u(t)\right] dt + A_p(\theta(t))x(t)dw(t),\; x(0) = x^0, \tag{6.52}$$

where

$$\begin{aligned}
&\bar{v}^*(t) := \bar{F}x(t) = \gamma^{-2}D^T\bar{Z}x(t),\\
&\bar{\boldsymbol{A}}(\theta(t)) := \sum_{k=1}^{M} \alpha_k \bar{\boldsymbol{A}}_k,\\
&B(\theta(t)) := \begin{bmatrix} B_1(\theta(t)) & \cdots & B_N(\theta(t)) \end{bmatrix}.
\end{aligned}$$

As the similar argument of the LQC problem in Lemma 6.1, if the following inequality holds, the closed-loop stochastic LPV system is mean square stable and it has a cost bound.

$$\begin{aligned}
\bar{P}\left(\bar{\boldsymbol{A}}(\theta) + B(\theta)\bar{K}\right) + \left(\bar{\boldsymbol{A}}(\theta) + B(\theta)\bar{K}\right)^T \bar{P}&\\
+ A_p^T(\theta)\bar{P}A_p(\theta) + Q + \bar{K}^T\boldsymbol{R}\bar{K} < 0,&
\end{aligned} \tag{6.53}$$

where $u(t) = u^*(t) = \bar{K}x(t),\; \bar{K} = \begin{bmatrix} \bar{K}_1 & \cdots & \bar{K}_N \end{bmatrix}$.

That is, the following equation holds.

$$J(x,\, x^0) < \mathbb{E}\left[x^T(0)\bar{P}x(0)\right]. \tag{6.54}$$

On the other hand, we have the following inequality by rearranging (6.53).

$$\begin{aligned}
&\sum_{k=1}^{M} \alpha_k^2 \left(\bar{P}(\bar{\boldsymbol{A}}_k + B_k\bar{K}) + (\bar{\boldsymbol{A}}_k + B_k\bar{K})^T\bar{P} + A_{pk}^T\bar{P}A_{pk} + \boldsymbol{Q} + \bar{K}^T\boldsymbol{R}\bar{K}\right)\\
&\quad \sum_{k=1}^{M-1}\sum_{l=k+1}^{M} \alpha_k\alpha_l \left(\bar{P}\bar{G}_{kl} + \bar{G}_{kl}^T\bar{P} + H_{kl} + 2\boldsymbol{Q} + 2\bar{K}^T\boldsymbol{R}\bar{K}\right) < 0,
\end{aligned} \tag{6.55}$$

where

$$\bar{G}_{kl} := \bar{\boldsymbol{A}}_k + B_k\bar{K} + \bar{\boldsymbol{A}}_l + B_l\bar{K}.$$

Therefore, as the sufficient conditions, if the following inequalities hold, inequality (6.55) is satisfied.

$$\bar{P}\left(\bar{\boldsymbol{A}}_k + B_k\bar{K}\right) + \left(\bar{\boldsymbol{A}}_k + B_k\bar{K}\right)^T \bar{P} + A_{pk}^T\bar{P}A_{pk} + Q + \bar{K}^T\boldsymbol{R}\bar{K} < 0, \tag{6.56a}$$

$$\begin{aligned}&\bar{P}\bar{G}_{kl} + \bar{G}_{kl}^T\bar{P} + H_{kl} + 2\boldsymbol{Q} + 2\bar{K}^T\boldsymbol{R}\bar{K}\\ &\quad \le \bar{P}\bar{G}_{kl} + \bar{G}_{kl}^T\bar{P} + 2A_{pkl}^TPA_{pkl} + 2\boldsymbol{Q} + 2\bar{K}^T\boldsymbol{R}\bar{K} < 0.\end{aligned} \tag{6.56b}$$

Thus, if inequalities (6.56) are satisfied, then inequality (6.55) holds. In other words, if inequalities (6.46a) and (6.46b) are satisfied, then (6.53) holds and this inequality implies the cost bound (6.54). □

6.6 Numerical Examples

In order to demonstrate the efficiency of our proposed strategies, an academic numerical example and a practical example based on air-path system of diesel engines are investigated.

6.6.1 Academic example

The system matrices are given as follows:

$$\begin{aligned}
&A_1 = \begin{bmatrix} -1.5 & 1 \\ -1 & -1.6 \end{bmatrix}, \quad A_{p1} = 0.1A_1,\\
&A_2 = \begin{bmatrix} -1.5 & 1 \\ -1 & -1.8 \end{bmatrix}, \quad A_{p2} = 0.1A_2,\\
&\alpha_1(t) = \sin t, \quad \alpha_2(t) = 1 - \sin t,\\
&B_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\\
&D = \begin{bmatrix} 0.22 \\ 0.2 \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0.4 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.3 \end{bmatrix},\\
&Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \quad R_1 = 5, \quad R_2 = 4.
\end{aligned}$$

The disturbance attenuation is chosen as $\gamma = 5$. The CCMIs (6.31) are by using algorithm of the previous subsection. The strategy-set (6.32) which attains the Pareto optimal solution with H_∞ constraint is given below.

$$K_1 = \begin{bmatrix} K_{11} \\ K_{21} \end{bmatrix} = \begin{bmatrix} 2.9247\text{e-}2 & -7.0586\text{e-}1 \\ 1.8279\text{e-}2 & -4.4116\text{e-}1 \end{bmatrix},$$

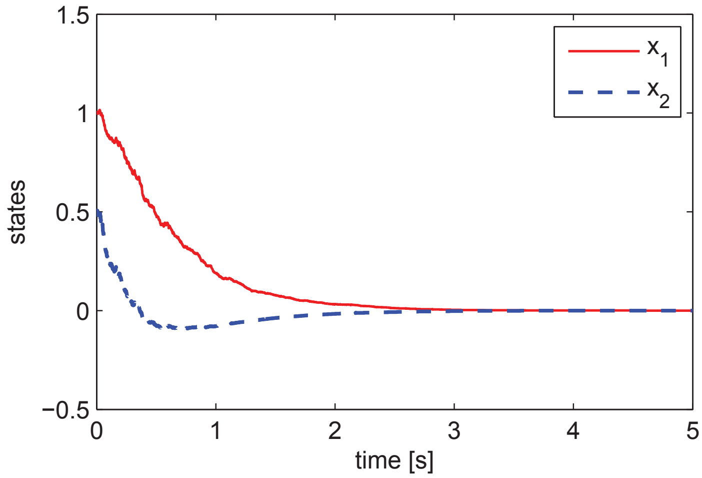


Fig. 6.1: Simulation results for the closed-loop system under the time-varying gain.

$$K_2 = \begin{bmatrix} K_{12} \\ K_{22} \end{bmatrix} = \begin{bmatrix} 2.9247\text{e-}2 & -7.0586\text{e-}1 \\ 1.8279\text{e-}2 & -4.4116\text{e-}1 \end{bmatrix},$$
$$F_\gamma = \begin{bmatrix} 1.0680\text{e-}3 & 8.9609\text{e-}4 \end{bmatrix}.$$

The proposed algorithm converges to the required solution with an accuracy of 1.0e-8 order only after 5 iterations. In order to verify the performance constraint condition by the value of the following ratio function is computed.

$$\|L\|^2_{[0,\ t_f]} = \frac{\mathbb{E}\left[\int_0^{t_f} \|z(t)\|^2 dt\right]}{\mathbb{E}\left[\int_0^{t_f} \|v(t)\|^2 dt\right]} = 2.8131\text{e-}2 < \gamma^2 = 25.$$

It can be observed that the value of the above ratio function (6.45) is small when $\gamma = 5$. Hence, the constraint condition is satisfied. Second, the time histories with $x(0) = \begin{bmatrix} 1 & 0.5 \end{bmatrix}$ are depicted from Fig. 6.1.

It should be noted that the disturbance is chosen as $v(t) = [1\ 1]\sin^2 t$. From Fig. 6.1, one can find that the asymptotic stability can be achieved. In other words, one can succeed in reducing the influence of the deterministic disturbance $v(t)$ by means of the designed Pareto optimal strategy-set.

6.6.2 Practical example (air-path system of diesel engines)

In order to demonstrate the effectiveness of the proposed method, we show results for the control problem on the air-path system of the diesel engine [Ku and Wu (2015)] with some trivial modifications. In [Ku and Wu (2015)], the gain-scheduled H_∞ control for stochastic

LPV system was set so that the disturbance does not affect the performance output more than performance index $\gamma > 0$. Although in [Ku and Wu (2015)] used state feedback control to stabilize the system (6.27), include the Pareto optimal control assuming a quadratic cost functional for each control input. The idea is based on [Zeng et al. (2017)], where a linear quadratic (LQ) controller design used to minimize the tracking errors of both exhaust gas re-circulation (EGR) mass flow rate and boost pressure through variable geometry turbocharger (VGT).

According to [Ku and Wu (2015)], $x = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) \end{bmatrix}^T$ and $u(t) = \begin{bmatrix} u_1(t) & u_2(t) \end{bmatrix}^T$ denote the state vector and control input, respectively. Furthermore, $x_1(t)$ denotes intake manifold pressure; $x_2(t)$ denotes exhaust manifold pressure; $x_3(t)$ denotes compressor air mass flow; $u_1(t)$ denotes EGR valve position; $u_2(t)$ denotes VGT vane position; and $v(t)$ is chosen as zero-mean white noise with variance one. Thus, the disturbed air-path of diesel engines can be described as stochastic system (6.27a), where

$$
\begin{aligned}
&A(\theta) = \begin{bmatrix} -12.6 & 8.2 & 0 \\ 1.01 + 4\theta_1 & -2.08 & 0 \\ 0.12 + 4.04\theta_1 & -0.37 - 0.44\theta_2 & -1.33 \end{bmatrix}, \\
&B_1 = \begin{bmatrix} 0 \\ -25.65 \\ -18.27 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 \\ 40.32 \\ 0 \end{bmatrix}, \; D = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \\
&A_p(\theta) = 0.01A(\theta), \; \theta = \theta(\theta_1, \; \theta_2),
\end{aligned}
$$

It should be noted that exact function of parameter $\theta(t)$ is generally unknown in this example and several cases should be simulated. According to the air-path system of diesel engines, the two scheduling parameters $\theta_1(t)$ and $\theta_2(t)$ vary within the following ranges [Liu et al. (2007)]:

$$
\begin{aligned}
\theta_1(t) &\in [-0.15, \; 0.15], \\
\theta_2(t) &\in [-0.84, \; 0.16].
\end{aligned}
$$

In this example, the following parameters are applied as a special case.

$$
\begin{aligned}
\theta_1(t) &= 0.15\sin(\omega t), \\
\theta_2(t) &= -0.34 - 0.5\sin(\omega t),
\end{aligned}
$$

where, ω is the frequency determined by the frequency response. After several trials, [Liu et al. (2007)] shows air mass flow step response and boost pressure step response for

$$\theta_1(t) = 0.15\sin(10t), \tag{6.57}$$

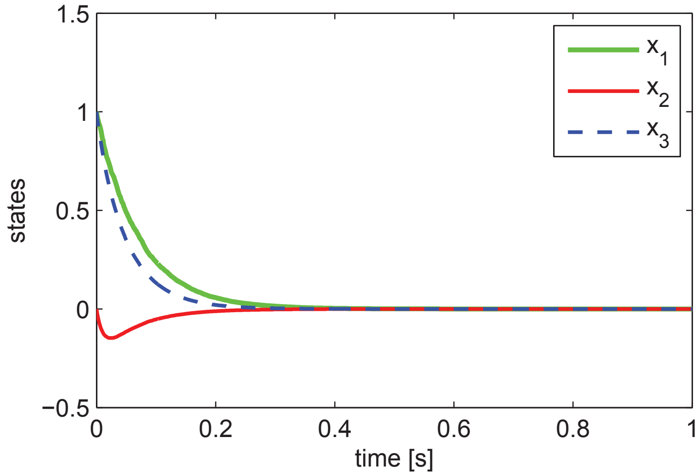


Fig. 6.2: Simulation results for LPV system under practical plant.

$$\theta_2(t) = -0.34 - 0.5\sin(10t), \tag{6.58}$$

LPV controller achieves much better performance. It can be seen that they change very quickly and cover the entire range of two scheduling variables, while rapid changes in scheduling variables have only a slight effect on the system response.

Since the number of time-varying parameters is 2, the number $M = 2^r = 4$. So, the parameter dependent coefficient matrices (6.2) can be decomposed as follows:

$$A_1 = \begin{bmatrix} -12.6 & 8.2 & 0 \\ 0.41 & -2.08 & 0 \\ -0.486 & -0.0004 & -1.33 \end{bmatrix},\ A_2 = \begin{bmatrix} -12.6 & 8.2 & 0 \\ 0.41 & -2.08 & 0 \\ -0.486 & -0.4404 & -1.33 \end{bmatrix},$$
$$A_3 = \begin{bmatrix} -12.6 & 8.2 & 0 \\ 1.61 & -2.08 & 0 \\ 0.726 & -0.0004 & -1.33 \end{bmatrix},\ A_4 = \begin{bmatrix} -12.6 & 8.2 & 0 \\ 1.61 & -2.08 & 0 \\ 0.726 & -0.4404 & -1.33 \end{bmatrix},$$
$$A_{pk} = 0.01A_k,\ \text{for } k = 1,\ 2,\ 3,\ 4.$$

Without loss of generality, we can assume $E_1 = E_2 = E_3 = E_4 = I_3$. The weight matrices for the cost functionals are assumed as $Q_1 = \mathbf{diag}(\,7 \quad 5 \quad 10\,)$, $Q2 = \mathbf{diag}(\,6 \quad 21 \quad 5\,)$, $R_1 = 9$, $R_2 = 9$. It should be noted that, the performance of the resulting closed-loop system can be adjusted by appropriately selecting the LQ weighting matrices [Zeng et al. (2017)]. The values for $\alpha_1 = a(1-b)$, $\alpha_2 = ab$, $\alpha_3 = (1-a)(1-b)$ and $\alpha_4 = (1-a)b$, where $a = (0.15-\theta_1)/0.3$ and $b = \theta_2 + 0.84$. The disturbance attenuation is chosen as $\gamma = 2$.

The CCMIs (6.31) are by using algorithm of the previous subsection. The strategy-set (6.32) which attains the Pareto optimal solution with H_∞ constraint is given below.

$$K_1 = \begin{bmatrix} K_{11} \\ K_{21} \end{bmatrix} = \begin{bmatrix} 1.2070\text{e-}1 & 7.4273\text{e-}1 & 1.1100\text{e+}0 \\ -2.5075\text{e-}1 & -1.5599\text{e+}0 & 5.5093\text{e-}1 \end{bmatrix},$$

$$K_2 = \begin{bmatrix} K_{12} \\ K_{22} \end{bmatrix} = \begin{bmatrix} 1.2070\text{e-1} & 7.4273\text{e-1} & 1.1100\text{e+0} \\ -2.5075\text{e-1} & -1.5599\text{e+0} & 5.5093\text{e-1} \end{bmatrix},$$
$$K_3 = \begin{bmatrix} K_{13} \\ K_{23} \end{bmatrix} = \begin{bmatrix} 1.2070\text{e-1} & 7.4273\text{e-1} & 1.1100\text{e+0} \\ -2.5075\text{e-1} & -1.5599\text{e+0} & 5.5093\text{e-1} \end{bmatrix},$$
$$K_4 = \begin{bmatrix} K_{14} \\ K_{24} \end{bmatrix} = \begin{bmatrix} 1.2070\text{e-1} & 7.4273\text{e-1} & 1.1100\text{e+0} \\ -2.5075\text{e-1} & -1.5599\text{e+0} & 5.5093\text{e-1} \end{bmatrix},$$
$$F_\gamma = \begin{bmatrix} -3.1441\text{e-4} & -2.5139\text{e-3} & 1.5667\text{e-2} \end{bmatrix}.$$

The proposed algorithm converges to the required solution with an accuracy of 1.0e-6 order after 6 iterations. In order to verify the performance constraint condition by the value of the following ratio function is computed.

$$\|L\|^2_{[0,\ t_f]} = \frac{\mathbb{E}\left[\int_0^{t_f} \|z(t)\|^2 dt\right]}{\mathbb{E}\left[\int_0^{t_f} \|v(t)\|^2 dt\right]} = 9.1315\text{e-2} < \gamma^2 = 4.$$

It can be observed that the value of the above ratio function is small when $\gamma = 2$. Hence, the constraint condition is satisfied. Second, the time histories with $x(0) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$ are depicted from Fig. 6.2. From Fig.6.2, it can be observed that the asymptotic stability can be achieved. In other words, one can succeed in reducing the influence of the deterministic disturbance $v(t)$ by means of the designed Pareto optimal strategy-set.

Remark 6.1. *Diesel engines generate Nitrogen Oxides (NO_x) emissions that are toxic and cause health problems. To reduce NO_x emissions, an effective means is to regulate transient exhaust gas re-circulation (EGR) using control strategies. On the other hand, by controlling the variable geometry turbine (VGT) vane position, the boost pressure is adjusted to save fuel efficiently. The VGT absorbs the waste-heat energy and recycling gas from the EGR to drive the compressor. On the other hand, when VGT vane is closed, the exhaust gas flowing into the EGR increases. So, there is a correlation between these two controls. However, [Zeng et al. (2017)] used LQ controllers to minimize the tracking errors of the EGR mass flow and boost pressure through VGT. Our idea is to associate these two controls to a cooperative game with a common objective and to adapt to Pareto optimality. The advantage of this an idea is that it not only minimizes the individual tracking errors, but also optimizes the overall performance of the diesel engine with disturbance attenuation under H_∞-constraint. Therefore, by applying the proposed scheme, it is possible to generate a more sophisticated, environmentally friendly fuel-efficient diesel engine.*

6.7 Conclusions

This chapter discusses the Pareto optimal strategy for stochastic LPV system with multiple decision makers. In the dynamic game of uncertain stochastic systems, multiple participants can be used for more realistic plants. The deterministic disturbances and their attenuation to stochastic LPV systems under the H_∞ constraint is the main attraction of this chapter. Problems involving deterministic disturbance must be attenuated at a given target called disturbance attenuation level $\gamma > 0$. This chapter can be seen as an extension of [Mukaidani (2017a)] in the sence that the fixed gain controller is also considered here. In this chapter, we design a method for Pareto optimal solution for multiple decision makers that satisfies the H_∞ norm condition.

Unlike the existing Pareto optimal strategy-set, the gain-scheduled controllers have been adopted for the first time. As a result, even though the deterministic time-varying parameters in the stochastic systems exist, a strategy-set can be designed. We redesigned the stochastic bounded real lemma [Ku and Wu (2015)] and the linear quadratic control [Rotondo (2015)] to find the solution. The Pareto optimal strategy-set can be found by solving a set of cross-coupling matrix inequalities (CCMIs). The modified stochastic bounded real lemma and linear quadratic control (LQC) for the stochastic LPV systems are reformulated by means of linear matrix inequalities (LMIs). The solvability conditions of the problem are established from cross-coupled matrix inequalities (CCMIs). Since these matrix inequalities are coupled, it is very complicated if an ordinary scheme such as Newton's method is applied. A numerical algorithm via the semidefinite programming problem (SDP) is developed to solve this problem.

The proof of convergence for the method based on the SDP (6.39) is not discussed. Moreover, the uniqueness of the solution was not proved. These problems will be addressed in future investigations. Academic and practical numerical examples show the feasibility of the proposed method. In order to demonstrate the real life application of the proposed method, we show results for the control problem on the air-path system of the diesel engine. Although we have not implemented the H_∞ constraint incentive Stackelberg game for stochastic LPV systems, this will be our future research. However, in our current research, the information structure is used as state feedback; the output feedback pattern will be investigated in our future studies.

Chapter 7

Conclusion

This thesis investigates the incentive Stackelberg game for discrete-time systems and continuous-time systems. Prior to this, the basic terminologies of the dynamic game are introduced. The motivation is to choose the incentive Stackelberg game to be an engineering application of a packet switch that works in a loop structure [Saksena and Cruz (1985)]. The above problem comes from a static game. However, this thesis studies only dynamic games.

For discrete time case, both deterministic and stochastic systems are investigated. Results based on finite and infinite time domains are shown in discrete time. However, stochastic systems are only considered in the case of continuous time. It should be noted that the generalized results given by stochastic investigation can also be applied to deterministic cases. To simplify the calculation, only the infinite time domain in the case of continuous time is emphasized. In most cases, the linear differential equation governed by Ito's differential equation is used in the theory of this research. This is a very common phenomenon in the field of control theory research and is simple to operate.

This thesis studies the most common linear quadratic (LQ) optimal control in game problems. To solve the LQ problem, stochastic dynamic programming (SDP) and stochastic maximum principle are deeply studied. Cooperative and non-cooperative game problems are solved based on the concepts of Pareto optimality and Nash equilibrium solutions, respectively. Several basic problems are completely solved and useful for the current research. The main task to solve the LQ problem is to find a matrix solution of algebraic Riccati equations. However, the Newton's method is very effective for fast convergence, the Lyapunov's iterative method is most popular for a simple built-in function 'lyap(·)'. Among the various styles for presenting results, figures for the trajectories of the states are the most attractive and reliable to ensure that the system is stable.

The deterministic disturbances and their attenuation to stochastic systems under the H_∞ constraint is the main attraction of this thesis. Problems involving deterministic disturbance must be attenuated at a given target called disturbance attenuation level $\gamma > 0$. Surprisingly, the concept of solving the disturbance reduction problem under the H_∞ constraint seems like a Nash equilibrium between the disturbance input and the control input.

In the incentive Stackelberg game, players are divided into two categories; the leader group and the follower group. For a single leader game, incentive Stackelberg strategy is an extensive idea in which the leader can achieve his/her team-optimal solution in a Stackelberg game. Multiple leaders and multiple followers have made the game more complex and challenging. In the leaders' and the followers' groups, the players are supposed to be non-cooperative; subsequently, the Nash equilibrium is investigated. Several novel theorems and lemmas are designed to study the incentive Stackelberg game problems. In this game, an incentive structure is developed in such a way that leaders achieve Nash equilibrium by attenuating the disturbance under H_∞ constraint. Simultaneously, followers achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance is considered. Interestingly, for all cases, some sets of cross-coupled matrix algebraic and Riccati equations can be derived to find the set of strategies. To solve those matrix equations, algorithms based on Lyapunov iterations are developed. In addition, several academic and real-life numerical examples have also been resolved to demonstrate the usefulness of our proposed scheme.

This thesis discusses the incentive mechanism of the Stackelberg game in detail, but it also gives a small description of the ordinary Stackelberg game. A detailed survey shows that over the past four decades, several studies have been conducted on the incentive Stackelberg game. However, the main objective of this research is to investigate the incentives Stackelberg strategy, preliminary research and synthesis of LPV systems for multiple decision makers. We aim to better understand to implement our current idea for LPV systems in the future. H_∞ constraint Pareto optimal strategy for stochastic linear parameter varying (LPV) systems with multiple decision makers is investigated. The modified stochastic bounded real lemma and LQ control for the stochastic LPV systems are reformulated by means of linear matrix inequalities (LMIs). To decide the strategy set of multiple decision makers, the Pareto optimal strategy is considered for each player and the H_∞ constraint is imposed. The solvability conditions of the problem are established from cross-coupled matrix inequalities (CCMIs).

The basics of LMIs are discussed as an appendix. However, the results and discussion on LMIs already exist. It gives an important idea in the formulation and solution of the control problems. The appendix discusses how to solve convex optimization problems using LMIs and special cases to solve systems and control theory problems. We consider the original problem from solution system and control theory. Although the appendix mainly covers system and control theory, there is a possibility to pose problems for convex optimization as well. The method described in that appendix has great practical value for control engineering. MATLAB LMI toolbox is an essential feature of the control theory research to solve LMI system.

In fact, through this thesis, stochastic games of multiple decision makers with disturbances open a new dimension of optimal control research. Several academic and practical numerical examples show the feasibility of the proposed method. Although we have not implemented the H_∞ constraint incentive Stackelberg game for stochastic LPV systems, this will be our future research. In our current research, the information structure is used as state feedback; the output feedback pattern will be investigated in our future studies.

Some preliminary results on static output feedback optimal control are given in Appendix B. The linear quadratic optimal cost control problem for static output feedback optimal control for stochastic Itô differential equations is considered. Several definitions, theorems, and lemmas are studied for future research. To solve the output feedback control problem, the Newton's algorithm and corresponding codes are already developed. A numerical example of fundamental problems has been solved. The problem is already formulated for future investigation.

In that problem, an infinite-horizon incentive Stackelberg game with multiple leaders and multiple followers will be investigated for a class of linear stochastic systems with H_∞ constraint. An incentive structure will be developed in such a way that leaders will achieve the Nash equilibrium by attenuating the disturbance under H_∞ constraint. Simultaneously, followers will achieve their Nash equilibrium ensuring the incentive Stackelberg strategies of the leaders while the worst-case disturbance will be considered. In that research, some cross-coupled stochastic algebraic Riccati equations (CCSAREs) and matrix algebraic equations (MAEs) will be derived for static output feedback case so that the incentive Stackelberg strategy-set can be found. Unlike current research, static output feedback control will be considered. However, [Mukaidani et al. (2018)] studied discrete-time linear stochastic systems with infinite time-domain incentives by means of Markov jump parameters and external disturbances through static output feedback. Multiple leader-follower problems with output feedback in continuous-time linear stochastic systems will be our future study. Moreover, in the future, we have plans to extend the proposed results to the output feedback by means of the state observer. Also, the information that the leader can utilize may be different from that the followers can utilize.

Appendix A

Linear Matrix Inequalities (LMIs)

The basic topic in this appendix is how to solve convex optimization problems using linear matrix inequalities (LMIs) and special cases to solve systems and control theory problems. Although this appendix mainly covers system and control theory, there is a possibility to pose problems for convex optimization as well. The method described in this appendix has great practical value for control engineering. LMI is a matrix inequality that is linear or affine within a set of matrix variables. Since they are convex constraints themselves, many existing software can efficiently and easily solve many convex objective functions and optimization problems of LMI constraints. This method has become very popular among control engineers in recent years. This is because various control issues can be formulated as LMI issues.

The LMI format is as follows:

$$G(x) = G_0 + x_1 G_1 + \cdots + x_m G_m > 0. \tag{A.1}$$

Where $x \in \mathbb{R}^m$ is the vector of decision variables, $G_0, G_1, \ldots, G_n$ are constant Symmetric matrices, that is

$$G_i = G_i^T, \quad i = 0, 1, \ldots, m. \tag{A.2}$$

The inequality used in equation (A.1) means the positive definiteness of $G(x)$, i.e., $u^T G(x) u > 0$, for all $u \in \mathbb{R}^n$.

Definition A.1. *[Rami and Zhou (2000)] Let $G_0, G_1, \ldots, G_m \in \mathscr{S}^n$ be constant symmetric matrices. Inequalities consisting of any combination of the following relations:*

$$G(x) := G_0 + \sum_{i=1}^{m} x_i G_i > 0 \tag{A.3a}$$

$$G(x) := G_0 + \sum_{i=1}^{m} x_i G_i \geq 0 \tag{A.3b}$$

are called LMI's for variable $x = (x_1, \ldots, x_m)^T \in \mathbb{R}^m$. If there is at least one $x \in \mathbb{R}^m$ that satisfies them, LMI is called feasible, and point x is called a feasible point.

LMIs is basically used in dynamical system of Lyapunov theory.

Theorem A.1. *The following differential equation,*

$$\dot{x}(t) = Ax(t) \tag{A.4}$$

is asymptotically stable iff there exists a matrix $P > 0$ such that the following the Lyapunov inequality holds:

$$A^T P + PA < 0. \tag{A.5}$$

Proof. Let us consider the Lyapunov candidate

$$V(t) = x^T(t)P(t)x(t), \tag{A.6}$$

where P is a symmetric positive definite matrix. Therefore,

$$\begin{aligned}\dot{V}(t) &= \dot{x}^T(t)Px(t) + x^T(t)P\dot{x}(t)\\ &= x^T(t)A^T Px(t) + x^T(t)PAx(t)\\ &= x^T(t)[A^T P + PA]x(t).\end{aligned} \tag{A.7}$$

According to Lyapunov stability theorem, for a Lyapunov function $V(t)$, the system (A.4) is asymptotically stable if and only if $V(t) < 0$. Comparing equations (A.6) and (A.7) we can obtain

$$A^T P + PA < 0. \tag{A.8}$$

Hence Theorem A.1 is proved. □

If we consider a matrix $Q = Q^T > 0$ such that

$$A^T P + PA = -Q,$$

the LMI turns to be a matrix algebraic equation.

A.1 Formation of LMIs

Many control problems can be expressed as LMI problems, but some of them cause nonlinear matrix inequalities. Specific techniques can be used to convert these nonlinear inequalities to the appropriate LMI format. Here, we will use appropriate examples to describe some of the techniques that are often used for control.

Let us consider a state feedback optimal control problem in which we have to find a state feedback gain matrix $K \in \mathbb{R}^{m\times n}$ and a symmetric positive definite matrix $P \in \mathbb{R}^{n\times n}$ such that the following inequality holds:

$$(A+BK)^T P + P(A+BK) < 0, \tag{A.9}$$

equivalently,

$$A^T P + PA + K^T B^T P + PBK < 0. \tag{A.10}$$

It should be noted that the matrices K and P are contained in the same product terms make the inequality nonlinear or bilinear. To make it linear, suppose that $X = P^{-1}$, which gives

$$XA^T + AX + XK^T B^T + BKX < 0. \tag{A.11}$$

This is also a matrix inequality containing a new variable X. However, the inequality is still nonlinear. Let us consider another new variable $L = KX$, which gives

$$XA^T + AX + L^T B^T + BL < 0. \tag{A.12}$$

This is a LMI feasibility problem with respect the variable $X > 0$ and $L \in \mathbb{R}^{m\times n}$. Solving this LMI problem, the feedback gain matrix K can be found from the relation $K = LX^{-1}$ and $P = X^{-1}$. This shows that by changing variable of a nonlinear matrix inequality problem into an LMI problem.

The Schur Complement can be used to transform nonlinear inequalities of convex type LMI problem.

Lemma A.1. *(Schur's lemma)[Rami and Zhou (2000)] Let matrices $Q = Q^T$, S, and $R = R^T > 0$ be given with appropriate dimensions. Then the following conditions are equivalent:*

$$1)\ Q - SR^{-1}S^T > 0;$$

$$2)\ \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} > 0;$$

$$3)\ \begin{bmatrix} R & S^T \\ S & Q \end{bmatrix} > 0.$$

For example, let us consider the following quadratic matrix inequality [Boyd et al. (1994)]:

$$A^T P + PA + PBR^{-1}B^T P + Q < 0, \tag{A.13}$$

where $A,\ B,\ Q = Q^T,\ R = R^T > 0$ are given matrices of appropriate sizes, and $P = P^T$ is the variable. It should be noted that this is a quadratic matrix inequality in the variable P. It can be expressed as the linear matrix inequality as follows:

$$\begin{bmatrix} -A^T P - PA - Q & PB \\ B^T P & R \end{bmatrix} > 0, \tag{A.14}$$

or,

$$\begin{bmatrix} R & B^T P \\ PB & -A^T P - PA - Q \end{bmatrix} > 0. \tag{A.15}$$

The MATLAB LMI toolbox provides some convenient functions for solving LMI problems. Now we present an example for solving a control problem by using MATLAB LMI toolbox.

Consider the following continuous-time stochastic linear quadratic optimal control problem:

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{A.16a}$$

$$J(x_0,\ u) := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left(x^T(t)Qx(t) + u^T(t)Ru(t)\right) dt\right], \tag{A.16b}$$

where $x(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^n)$ is the state vector; $u(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ is the control input; $w(t) \in \mathbb{R}$ is a one-dimensional wiener process; A, B, A_p, $Q = Q^T \geq 0$, $R = R^T > 0$ are the coefficient matrices of suitable dimensions. For the stochastic optimal control problem (A.16), suppose that the following stochastic ARE has the solution $P^T = P > 0$:

$$PA + A^T P + Q - PBR^{-1}B^T P + A_p^T PA_p = 0. \tag{A.17}$$

then the optimal control problem admits a state feedback solution,

$$u^*(t) = Kx(t) = -R^{-1}B^T Px(t). \tag{A.18}$$

LMI associated to (A.17) can be written as:

$$\left[\begin{array}{c|c} A^T P + PA + Q + A_p^T PA_p & PB \\ \hline B^T P & R \end{array}\right] \geq 0, \tag{A.19}$$

with respect to the symmetric matrix variable P. It should be noted that Definition A.1 can be applied to LMI (A.19) by a simple transformation. Consider $P_1,\ \ldots,\ P_{n(n+1)/2}$ be any basis of $\mathscr{S}_n$. The variable matrix P can be written as

$$P := \sum_{i=1}^{n(n+1)/2} x_i P_i.$$

Considering $(x_1, \dots, x_{n(n+1)/2})$ as a new variable, we can see that (A.19) is followed by Definition A.1.

A.2 Some standard LMI problems

Given an LMI $F(x) > 0$, the corresponding LMI problem is to find x_{feas} so that $F(x_{feas}) > 0$ or determine that the LMI is infeasible. For example, let us consider the following the simultaneous Lyapunov stability problem:

$$P > 0, \quad A_i^T P + PA_i < 0, \quad i = 1, \dots, L, \tag{A.20}$$

where $A_i \in \mathbb{R}^{n \times n}$, $i = 1, \dots, L$. It is needed to find P that satisfies the LMI (A.20) or to determine such P does not exist.

For the linear stochastic system

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw(t), \quad x(0) = x^0, \tag{A.21}$$

the following theorem will verify the stability with some equivalence conditions.

Theorem A.2. *[Rami and Zhou (2000)]*

1. *System* (A.21) *is mean-square stable.*

2. *There is a matrix K and $X = X^T > 0$ such that*

$$(A+BK)^T X + X(A+BK) + A_p^T X A_p < 0. \tag{A.22}$$

   The state feedback $u(t) = Kx(t)$ is stable, in this case.

3. *There is a matrix K and $X = X^T > 0$ such that*

$$(A+BK)X + X(A+BK)^T + A_p X A_p^T < 0. \tag{A.23}$$

   The state feedback $u(t) = Kx(t)$ is stable, in this case.

4. *There exists a matrix K such that for any matrix Y there is a unique solution X to the following equation*

$$(A+BK)^T X + X(A+BK) + A_p^T X A_p + Y = 0. \tag{A.24}$$

   Furhtermore, if $Y > 0$ (respectively, $Y \geq 0$), then $X > 0$ (respectively, $X \geq 0$). Moreover, the state feedback $u(t) = Kx(t)$ is stable, in this case.

5. *There exists a matrix K such that for any matrix Y there is a unique solution X to the following equation*

$$(A+BK)X+X(A+BK)^T+A_pXA_p^T+Y=0. \tag{A.25}$$

Furhtermore, if $Y>0$ (respectively, $Y\geq 0$), then $X>0$ (respectively, $X\geq 0$). Moreover, the state feedback $u(t)=Kx(t)$ is stable, in this case.

6. *There is a matrix Y and a symmetric matrix X such that*

$$\left[\begin{array}{c|c} AX+XA^T+BY+Y^TB^T & A_pX \\ \hline XA_p^T & -X \end{array}\right]<0. \tag{A.26}$$

In this case, the state feedback $u(t)=YX^{-1}x(t)$ is stable.

Proof. For any matris $K\in\mathbb{R}^{n_u\times n}$, define the following operator

$$\phi:\mathcal{S}^n\to\mathcal{S}^n, \tag{A.27}$$

by

$$\phi(X)=(A+BK)X+X(A+BK)^T+A_pXA_p^T. \tag{A.28}$$

If $x(\cdot)$ satisfies the following state feedback equation

$$dx(t)=[A+BK]x(t)dt+A_px(t)dw(t),\quad x(0)=x^0, \tag{A.29}$$

where, K is a feedback gain, then applying Itô's formula, the matrix $X(t)=\mathbb{E}[x^T(t)x(t)]$ satisfies the differential matrix system $\dot{X}(t)=\phi(X(t))$. Applying the results given in [Ghaoui and Rami (1996)], we have the equivalence between the mean-square stabilizability and each of the assertions 2–5. Furthermore, with $Y=KX$ and $X>0$, (A.23) is equivalent to

$$AX+XA^T+BY+Y^TB^T+A_pXA_p^T<0. \tag{A.30}$$

All other equivalence relations can be established by applying the Schur's lemma, as shown in Lemma A.1 here. □

Appendix B

Future Research

B.1 Static output feedback optimal control

In this section, we consider the problem of linear quadratic optimal cost control via static output feedback optimal control. The technique can be described by the stochastic differential equation of state as follows:

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw, \quad x(0) = x_0, \tag{B.1a}$$

$$y(t) = Cx(t). \tag{B.1b}$$

where, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input, $w(t) \in \mathbb{R}$ is a one-dimensional wiener process, $y(t) \in \mathbb{R}^\ell$ is the output, $A,\ A_p \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{\ell\times n}$ are given coefficient matrices.

Definition B.1. *[Dragan et al. (2006)] The system* (B.1) *or* $(A,\ B,\ A_p)$ *is called stochastic stabilizable (in mean-square sense), if there exists a output feedback control* $u(t) = Ky(t) = KCx(t)$ *with* K *being a constant matrix, such that the closed-loop system*

$$dx(t) = [A + BKC]x(t)dt + A_p x(t)dw(t), \quad x(0) = x_0, \tag{B.2}$$

is asymptotically mean-square stable, i.e.

$$\lim_{t\to\infty} \mathbb{E}[\|x(t)\|^2] = 0. \tag{B.3}$$

Remark B.1. *Under the condition* $B \equiv 0$*,* $(A,\ A_p)$ *is called stable, if equation* (B.3) *holds.*

Definition B.2. *[Chen and Zhang (2004)] The state-measurement system* (B.1) *or* $(A,\ A_p \mid C)$ *is called stochastically detectable, if there exists a constant matrix* X *such that* $(A + XC,\ A_p)$ *is asymptotically mean-square stable.*

Lemma B.1. *[Chen and Zhang (2004)] If* $(A,\ C \mid \sqrt{Q})$ *is stochastic detectable, then the autonomous system* $(A,\ A_p)$ *is stable if and only if the following stochastic algebraic Lyapunov equation (SALE)* (B.4) *has a unique solution* $P \geq 0$*:*

$$A^T P + PA + A_p^T PA_p + Q = 0, \tag{B.4}$$

with cost functional

$$J(x_0) = \mathbb{E}\left[\int_0^\infty x^T(t)Qx(t)dt\right] = x_0^T Px_0, \tag{B.5}$$

where $Q = Q^T \geq 0$.

Proof. To prove the first part of Lemma B.1, the same procedure as Lemma 2.2 can be applied. To prove the second part

$$\begin{aligned}\mathbb{E}\left[\int_0^t x^T(s)Qx(s)ds\right] &= -\mathbb{E}\left[\int_0^t x^T(s)[A^T P + PA + A_p^T PA_p]x(s)ds\right]\\ &= -\mathbb{E}\left[\int_0^t x^T(s)\dot{P}x(s)ds\right] \quad \text{[Itô's formula for finite-horizon.]}\\ &= x_0^T Px_0 - \mathbb{E}[x^T(t)Px(t)] \to x_0^T Px_0, \quad \text{when } t \to \infty.\end{aligned}$$

□

Now, consider the stochastic optimal control problem (B.1) with the following cost functional:

$$J(u,\ x_0) = \mathbb{E}\left[\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)]dt\right], \tag{B.6}$$

where $Q = Q^T \geq 0$ and $R = R^T > 0$. Suppose that there exists an optimal state feedback control

$$u^*(t) = KCx(t), \tag{B.7}$$

where K is the feedback gain matrix of the static output feedback control problem (B.1) with $(A,\ A_p \mid C)$ is stochastically detectable.

Applying Lemma B.1, there exists a positive semi-definite matrix P, which is the solution of the following algebraic Riccati equation:

$$\mathscr{G}_1(P,\ K) := (A + BKC)^T P + P(A + BKC) + A_p^T PA_p + C^T K^T RKC + Q = 0. \tag{B.8}$$

To find the feedback gain K, consider the following Lagrangian:

$$\mathscr{L}(P,\ V,\ K) = \mathbf{Tr}[x_0 x_0^T P] + \mathbf{Tr}\left[\mathscr{G}_1(P,\ K)V\right], \tag{B.9}$$

where $V = V^T$; $V \in \mathbb{R}^{n\times n}$ is the Lagrange multiplier. On the other hand, $P \in \mathbb{R}^{n\times n}$ and $K \in \mathbb{R}^{m\times \ell}$ are the optimization variables. Taking partial derivatives of (B.9) with respect to V, P and K we can find the following results:

$$\frac{\partial \mathscr{L}(P,\ V,\ K)}{\partial V} := [(A+BKC)^T P + P(A+BKC) + A_p^T P A_p + C^T K^T RKC + Q]^T, \tag{B.10a}$$

$$\frac{\partial \mathscr{L}(P,\ V,\ K)}{\partial P} := (A+BKC)V + V(A+BKC)^T + A_p V A_p^T + (x_0^T x_0)I, \tag{B.10b}$$

$$\frac{\partial \mathscr{L}(P,\ V,\ K)}{\partial K} := 2B^T PVC^T + 2RKCVC^T. \tag{B.10c}$$

By setting $\partial\mathscr{L}/\partial V$, $\partial\mathscr{L}/\partial P$, and $\partial\mathscr{L}/\partial K$ equal to zero, we can find

$$(A+BKC)^T P + P(A+BKC) + A_p^T P A_p + C^T K^T RKC + Q = 0, \tag{B.11a}$$

$$(A+BKC)V + V(A+BKC)^T + A_p V A_p^T + (x_0^T x_0)I = 0, \tag{B.11b}$$

$$K = -R^{-1} B^T PVC^T (CVC^T)^{-1}. \tag{B.11c}$$

By solving the resulting equations of (B.11) at the same time, the optimal solution for K can be found. Newton's algorithm is proposed to solve the system (B.11).

B.1.1 Newton's algorithm

Inputs:

Let $P = P^{(0)}$, $V = V^{(0)}$ and $K = K^{(0)}$ be the given initial matrices; $ITER$ is the maximum number of iterations; TOL is the tolerance of convergence.

Outputs:

Solution matrices P, V and K.

Step 1 Let us consider the following nonlinear matrix functions.

$$\mathscr{G}_1(P,V,K) = (A+BKC)^T P + P(A+BKC) + A_p^T P A_p + C^T K^T RKC + Q, \tag{B.12a}$$

$$\mathscr{G}_2(P,V,K) = (A+BKC)V + V(A+BKC)^T + A_p V A_p^T + (x_0^T x_0)I, \tag{B.12b}$$

$$\mathscr{G}_3(P,V,K) = RKCVC^T + B^T PVC^T, \tag{B.12c}$$

Step 2 For $k = 1,\ 2,\cdots,\ ITER$ do **Step 3** to **Step 4**.

Step 3 Calculate the following newtons formula:

$$\mathrm{vec}X^{(k+1)} = \mathrm{vec}X^{(k)} - \left[\frac{\partial \mathrm{vec}\mathscr{G}(X)}{\partial(\mathrm{vec}X)^T}\bigg|_{X=X^{(k)}}\right]^{-1} \mathrm{vec}\mathscr{G}(X^{(k)}), \tag{B.13}$$

where

$$\mathrm{vec}\mathscr{G}(X) = \begin{bmatrix} \mathrm{vec}\mathscr{G}_1(P,V,K) \\ \mathrm{vec}\mathscr{G}_2(P,V,K) \\ \mathrm{vec}\mathscr{G}_3(P,V,K) \end{bmatrix} \text{ and } X = \begin{bmatrix} \mathrm{vec}P \\ \mathrm{vec}V \\ \mathrm{vec}K \end{bmatrix}.$$

Step 4 If $\|X^{(k+1)} - X^{(k)}\| < TOL$, **stop**.

Step 5: Output

Step 6: End

It should be noted that to compute (B.13), Jacobian $\frac{\partial \mathrm{vec}\mathscr{G}(X)}{\partial (\mathrm{vec}X)^T}$ can be defined as follows:

$$\frac{\partial \mathrm{vec}\mathscr{G}(X)}{\partial (\mathrm{vec}X)^T} = \begin{bmatrix} \frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}P)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}V)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}K)^T} \\ \frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}P)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}V)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}K)^T} \\ \frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}P)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}V)^T} & \frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}K)^T} \end{bmatrix},$$

where

$$\frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}P)^T} = (A+BKC)^T \otimes I_n + I_n \otimes (A+BKC)^T + A_p^T \otimes A_p^T, \quad \frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}V)^T} = 0,$$

$$\frac{\partial \mathrm{vec}\mathscr{G}_1(X)}{\partial (\mathrm{vec}K)^T} = C^T \otimes PB + (PB \otimes C^T)U_{m\ell} + C^T \otimes C^T K^T R + (C^T K^T R \otimes C^T)U_{m\ell},$$

$$\frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}P)^T} = 0, \quad \frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}V)^T} = I_n \otimes (A+BKC) + (A+BKC) \otimes I_n + A_p \otimes A_p,$$

$$\frac{\partial \mathrm{vec}\mathscr{G}_2(X)}{\partial (\mathrm{vec}K)^T} = VC^T \otimes B + (B \otimes VC^T)U_{m\ell},$$

$$\frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}P)^T} = CV \otimes B^T, \ \frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}V)^T} = C \otimes (RKC + B^T P), \ \frac{\partial \mathrm{vec}\mathscr{G}_3(X)}{\partial (\mathrm{vec}K)^T} = CVC^T \otimes R,$$

$U_{m\ell}$ denotes a permutation matrix in Kronecker matrix sense [Henderson and Searle (1981)] such that $U_{m\ell}\mathrm{vec}K = \mathrm{vec}K^T$, $K \in \mathbb{R}^{m\times\ell}$.

B.1.2 Numerical example

Let us consider the following system matrices:

$$A = \begin{bmatrix} -2.98 & 0.93 & 0 & -0.034 \\ -0.99 & -0.21 & 0.035 & -0.0011 \\ 0 & 0 & 0 & 1 \\ 0.39 & -5.555 & 0 & -1.89 \end{bmatrix}, \ B = \begin{bmatrix} -0.032 \\ 0 \\ 0 \\ -1.6 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ A_p = 0.1A, \ R = 1, \ Q = I_4, \ x_0 = \begin{bmatrix} 1 & 0.5 & -0.5 & -1 \end{bmatrix}^T.$$

Applying Newton's Algorithm B.1.1 through MATLAB simulations that meet the appropriate initial conditions of Newton-Kantorovich's theorem, the Algorithm B.1.1 provides the following results:

$$P = \begin{bmatrix} 1.0310 & -2.7190 & -1.9989 & -0.3816 \\ -2.7190 & 11.4489 & 2.7210 & -0.7817 \\ -1.9989 & 2.7210 & -8.6677 & -1.4500 \\ -0.3816 & -0.7817 & -1.4500 & 0.4378 \end{bmatrix},$$

$$V = \begin{bmatrix} 0.5193 & 0.7023 & -2.0968 & -7.6932 \\ 0.7023 & 2.8654 & 1.1328 & -1.9652 \\ -2.0968 & 1.1328 & 1.3150 & -1.2389 \\ -7.6932 & -1.9652 & -1.2389 & -2.2189 \end{bmatrix},$$

$$K = \begin{bmatrix} -4.5034 & -1.3158 \end{bmatrix}, \ \mathbf{Tr}[x_0 x_0^T P] = 0.1786.$$

k	$\|\mathscr{G}(X^{(k)})\|$
0	3.7675
1	3.2866×10^{-2}
2	5.0958×10^{-4}
3	2.9617×10^{-10}
4	8.052×10^{-15}

Table B.1: Error in each iteration.

It should be noted that algorithm B.1.1 converges to the exact solution with an accuracy of $\|\mathscr{G}(X^{(k)})\| < 10^{-14}$ only after four iterations. From Table B.1, it can be observed that Newton's method attains quadratic convergence under the appropriate initial conditions.

B.1.3 Future investigation

In this section, I would like to formulate a problem of H_∞-constrained multiple leaders, multiple followers incentive Stackelberg game with static output feedback. However, the results and discussions on this issue will be future investigations. Consider a linear stochastic system governed by the Itô differential equation defined by

$$dx(t) = \Bigg[Ax(t) + \sum_{i=1}^{M} [B_{Li1} u_{Li1}(t) + \ \dots \ + B_{LiN} u_{LiM}(t)] + \sum_{j=1}^{N} [B_{Fj1} u_{Fj1}(t) + \ \dots \ + B_{FjM} u_{FjM}(t)] + Dv(t)\Bigg] dt + A_p x(t) dw(t), \ \ x(0) = x^0, \tag{B.14a}$$

$$z(t) = \mathbf{col}\begin{bmatrix} Cx(t) & u_{c1}(t) & \dots & u_{cM}(t) \end{bmatrix}, \tag{B.14b}$$

$$y(t) = Ex(t), \tag{B.14c}$$

$$u_{ci}(t) = \mathbf{col}\left[u_{Li1}(t)\ \dots\ u_{LiN}(t) \quad u_{F1i}(t)\ \dots\ u_{FNi}(t)\right], \tag{B.14d}$$

where $x(t) \in \mathbb{R}^n$ represents the state vector; $z(t) \in \mathbb{R}^{n_z}$ represents the controlled output; $y(t) \in \mathbb{R}^{n_y}$ represents the measured output; $u_{Lij}(t) \in \mathbb{R}^{m_{Lij}}$ represents the leader L_i's control input for the follower F_j, $i = 1,\ \dots\ ,M,\ j = 1,\ \dots\ ,N$; $u_{Fji}(t) \in \mathbb{R}^{m_{Fji}}$ represents the follower F_j's control input according to the leader L_i, $i = 1,\ \dots\ ,M,\ j = 1,\ \dots\ ,N$; $v(t) \in \mathbb{R}^{m_v}$ represents the exogenous disturbance signal; $w(t) \in \mathbb{R}$ represents a one-dimensional standard Wiener process defined in the filtered probability space $(\Omega,\ \mathscr{F},\ \mathscr{P},\ \mathscr{F}_t)$ with $\mathscr{F}_t = \sigma\{w(s)\ :\ 0 \le s \le t\}$ [Chen and Zhang (2004)]. Cost functionals of the leaders L_i, $i = 1,\ \dots\ ,M$, are given by

$$\begin{aligned} &J_{Li}\left(u_{Li1},\ \dots\ ,u_{LiN},\ u_{F1i},\ \dots\ ,u_{FNi},\ v\right) \\ &\quad := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)Q_{Li}x(t) + \sum_{j=1}^N \left[u_{Lij}^T(t)R_{Lij}u_{Lij}(t) + u_{Fji}^T(t)R_{LFji}u_{Fji}(t)\right]\right\}dt\right], \end{aligned} \tag{B.15}$$

where $Q_{Li} = Q_{Li}^T \ge 0$, $R_{Lij} = R_{Lij}^T > 0$, $R_{LFji} = R_{LFji}^T \ge 0$, $i = 1,\ \dots\ ,M,\ j = 1,\ \dots\ ,N$. Cost functionals of the followers F_i, $i = 1,\ \dots\ ,N$ are given by

$$\begin{aligned} &J_{Fi}\left(u_{L1i},\ \dots\ ,u_{LMi},\ u_{Fi1},\ \dots\ ,u_{FiM},\ v\right) \\ &\quad := \frac{1}{2}\mathbb{E}\left[\int_0^\infty \left\{x^T(t)Q_{Fi}x(t) + \sum_{j=1}^M \left[u_{Lji}^T(t)R_{FLji}u_{Lji}(t) + u_{Fij}^T(t)R_{Fij}u_{Fij}(t)\right]\right\}dt\right], \end{aligned} \tag{B.16}$$

where $Q_{Fj} = Q_{Fj}^T \ge 0$, $R_{Fij} = R_{Fij}^T > 0$ and $R_{FLji} = R_{FLji}^T \ge 0$, $i = 1,\ \dots\ ,N,\ j = 1,\ \dots\ ,M$. For a two-level incentive Stackelberg game, leaders announce the following incentive strategy to the followers in ahead of time:

$$u_{Lij}(t) = \Lambda_{ji}x(t) + \Xi_{ji}u_{Fji}(t), \quad i = 1,\ \dots\ ,M \quad j = 1,\ \dots\ ,N, \tag{B.17}$$

where the parameters Λ_{ji} and Ξ_{ji} are to be determined associated with the Nash equilibrium strategies $u_{Fji}(t)$ of the followers for $i,\ \dots\ ,M\ j = 1,\ \dots\ ,N$. In this game, leaders will achieve a Nash equilibrium solution attenuating the external disturbance with H_∞ constraint. The infinite-horizon multi-leader-follower incentive Stackelberg games for linear stochastic systems with H_∞ constraint can be formulated as follows.

For any disturbance attenuation level $\gamma > 0$, to find, if possible, the static output feedback strategy $u_{Lij}^*(t) = K_{cij}y(t)$ and $u_{Fji}^*(t) = K_{Fji}y(t)$ such that

(i) the trajectory of the closed-loop system (B.14) satisfies the Nash equilibrium conditions (B.18a) of the leaders with H_∞ constraint condition (B.18b):

$$J_{Li}(u_{c1}^*, \dots, u_{cM}^*, v^*) \le J_{Li}(\gamma_{-i}^*(u_{ci})), v^*), \tag{B.18a}$$

$$0 \le J_v(u_{c1}^*, \dots, u_{cM}^*, v^*) \le J_v(u_{c1}^*, \dots, u_{cM}^*, v), \tag{B.18b}$$

where $i = 1, \dots, M$,

$$J_v(u_{c1}, \dots, u_{cM}, v) = \mathbb{E}\left[\int_0^\infty \left\{\gamma^2 \|v(t)\|^2 - \|z(t)\|^2\right\} dt\right], \tag{B.19}$$

$$\|z(t)\|^2 = x^T(t)C^T Cx(t) + \sum_{i=1}^{M} u_{ci}^T(t)u_{ci}(t), \tag{B.20}$$

$\forall\, v(t) \ne 0 \in \mathbb{R}^{m_v}$,
(ii) with a worst-case disturbance $v^*(t) \in \mathbb{R}^{m_v}$, follower's decision $u_{Fji}^*(t)) \in \mathbb{R}^{m_{Fji}}$; $i = 1, \dots, M,\ j = 1, \dots, N$ satisfies the following Nash equilibrium conditions:

$$J_{Fj}(u_{F1}^*, \dots, u_{FN}^*, v^*) \le J_{Fj}(\gamma_{-j}^*(\hat{u}_{Fj})), v^*), \tag{B.21}$$

where

$$\hat{u}_{Fj}(t) = \mathbf{col}\left[u_{Fj1}(t) \quad \dots \quad u_{FjM}(t)\right],\ j = 1, \dots, N.$$

It should be noted that $u_{Lij}(t)$ depend on $u_{Fji}(t)$ according to the incentive Stackelberg structures assumed in (B.17).

Remark B.2. *Our research plan is to extend the above results of the static output feedback incentive Stackelberg game for Markov jump linear stochastic system with disturbance. However, [Mukaidani et al. (2018)] investigated discrete-time linear stochastic systems with infinite time-domain incentives by means of Markov jump parameters and external disturbances through static output feedback. To the best of our knowledge, multiple leader-follower problems with output feedback in continuous-time Markov jump linear stochastic systems have not been studied yet.*

B.2 Observer-based output feedback control

To extend the proposed results to the output feedback by means of the state observer, some preliminary results have been studied.

Definition B.3. *[Chen and Zhang (2004)] Consider the following autonomous stochastic system with measurement equation.*

$$dx(t) = Ax(t)dt + A_p x(t)dw(t), \quad x(0) = x_0, \tag{B.22a}$$

$$y(t) = Cx(t), \tag{B.22b}$$

where $x(t) \in \mathbb{R}^n$ is the state vector; $y(t) \in \mathbb{R}^\ell$ is the measurement output; $w(t) \in \mathbb{R}$ is a one-dimensional wiener process; $A,\ A_p \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{\ell\times n}$ are the coefficient matrices. If $y(t) \equiv 0$, $\forall\, t \geq 0$ implies $x_0 = 0$, $(A,\ A_p \mid C)$ is called exactly observable.

To check the exact observability for the system (B.22) we can find the following obsevability matrix [Zhang and Chen (2004)]:

$$\mathscr{O}_s = \begin{bmatrix} C \\ CA \\ CA_p \\ CA_pA \\ CAA_p \\ CA^2 \\ CA_p^2 \\ \vdots \end{bmatrix}.$$

Then, $(A,\ A_p|C)$ is exactly observable iff $\operatorname{rank}(\mathscr{O}_s) = n$.

B.2.1 Basic problem

Here we consider the problem of linear quadratic optimal cost control via observer-based control design. The technique can be described by the stochastic differential equation of state as follows:

$$dx(t) = [Ax(t) + Bu(t)]dt + A_p x(t)dw, \quad x(0) = x_0, \tag{B.23a}$$

$$y(t) = Cx(t). \tag{B.23b}$$

where $x(t) \in \mathbb{R}^n$ is the state vector; $u(t) \in \mathbb{R}^m$ is the control input; $w(t) \in \mathbb{R}$ is a one-dimensional wiener process; $y(t) \in \mathbb{R}^\ell$ is the measurement output; $A,\ A_p \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{\ell\times n}$ are coefficient matrices.

Consider stochastic LQ control with the following cost functional:

$$J = \mathbb{E}\left[\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)]dt\right], \tag{B.24}$$

where $Q = Q^T \geq 0$ and $R = R^T > 0$.

B.2.2 Preliminary results

If the system state of (B.23) is not fully accessible, the state feedback controller may be disabled. This is the motivation to propose an output feedback controller; the controller will be an observer-based controller in the following form [Gao and Shi (2013)]:

$$d\hat{x}(t) = [A\hat{x}(t) + Bu(t)]dt + G[\hat{y}(t) - y(t)]dt, \quad \hat{x}(0) = x_0, \tag{B.25a}$$

$$\hat{y}(t) = C\hat{x}(t), \tag{B.25b}$$

$$u(t) = K\hat{x}(t), \tag{B.25c}$$

where $\hat{x}(t) \in \mathbb{R}^n$ is the estimation of $x(t)$; $\hat{y}(t) \in \mathbb{R}^\ell$ is the estimation of $y(t)$; $G \in \mathbb{R}^{n\times\ell}$ and $K \in \mathbb{R}^{m\times n}$ the observer gain and control gain, respectively.

So, by using (B.23) and (B.25), the following closed loop systems can be obtained:

$$dx(t) = [Ax(t) + BK\hat{x}(t)]dt + A_p x(t)dw, \quad x(0) = x_0, \tag{B.26a}$$

$$d\hat{x}(t) = [A\hat{x}(t) + BK\hat{x}(t)]dt + GC[\hat{x}(t) - x(t)]dt, \quad \hat{x}(0) = x_0. \tag{B.26b}$$

The cost functional of the state dynamics and the estimated state dynamics can be written as:

$$J = \mathbb{E}\left[\int_0^\infty [x^T(t)Qx(t) + \hat{x}^T(t)K^T RK\hat{x}(t)]dt\right], \tag{B.27}$$

where $Q = Q^T \geq 0$ and $R = R^T > 0$. Let $e(t) = x(t) - \hat{x}(t)$ be the error. Then from (B.26), we can find the error dynamics as

$$dx(t) = [Ax(t) + BK\hat{x}(t)]dt + A_p x(t)dw, \quad x(0) = x_0, \tag{B.28a}$$

$$de(t) = [Ae(t) + GCe(t)]dt + A_p x(t)dw, \quad e(0) = \mathbf{0} \in \mathbb{R}^n. \tag{B.28b}$$

The cost functional of the state and the error dynamics can be written as:

$$J = \mathbb{E}\left[\int_0^\infty [x^T(t)(Q + K^T RK)x(t) - 2x^T(t)K^T RKe(t) + e^T(t)K^T RKe(t)]dt\right], \tag{B.29}$$

where $Q = Q^T \geq 0$ and $R = R^T > 0$.

Let $\bar{x}(t) = \mathbf{col}[x(t) \quad e(t)]$. Therefore, form (B.26), a closed loop system can be obtained as follow:

$$d\bar{x}(t) = \bar{A}\bar{x}(t)dt + \bar{A}_p\bar{x}(t)dw(t), \quad \bar{x}(0) = \bar{x}_0, \tag{B.30}$$

where

$$\bar{A} = \begin{bmatrix} A + BK & -BK \\ 0 & A + GC \end{bmatrix}, \quad \bar{A}_p = \begin{bmatrix} A_p & 0 \\ A_p & 0 \end{bmatrix}, \quad \bar{x}_0 = \begin{bmatrix} x_0 \\ \mathbf{0} \end{bmatrix}.$$

The cost functional in this case is

$$J = \mathbb{E}\left[\int_0^{\infty} \bar{x}^T(t)\bar{Q}\bar{x}(t)dt\right], \tag{B.31}$$

where

$$\bar{Q} := \begin{bmatrix} Q + K^T RK & -K^T RK \\ -K^T RK & K^T RK \end{bmatrix}. \tag{B.32}$$

Definition B.4. *[Dragan et al. (2006)] The closed-loop system* (B.30) *or* $(\bar{A},\ \bar{A}_p)$ *is called asymptotically mean-square stable if*

$$\lim_{t\to\infty} \mathbb{E}[\|\bar{x}(t)\|^2] = 0. \tag{B.33}$$

Lemma B.2. *[Gao and Shi (2013)] If there exists a symmetric positive definite matrix* $\bar{Z}$ *such that*

$$\bar{A}^T\bar{Z} + \bar{Z}\bar{A} + \bar{A}_p^T\bar{Z}\bar{A}_p + \bar{Q} < 0, \tag{B.34}$$

then the stochastic system (B.30) *is mean-square stable.*

Theorem B.1. *[Gao and Shi (2013)] The closed-loop system* (B.30) *is mean-square stable if there are matrices* K, G, $X = X^T > 0$ *and* $Y = Y^T > 0$, *and a positive scalar* α *such that the following inequalities hold:*

$$\begin{bmatrix} \Lambda_1 + \alpha A_p^T Y A_p & -X(BK) - K^T RK \\ -(BK)^T X - K^T RK & \alpha\Lambda_2 + K^T RK \end{bmatrix} < 0, \tag{B.35}$$

where

$$\Lambda_1 = (A+BK)^T X + X(A+BK) + A_p^T X A_p + Q + K^T RK, \tag{B.36}$$

$$\Lambda_2 = (A+GC)^T Y + Y(A+GC). \tag{B.37}$$

Proof. Let $\bar{Z} = \begin{bmatrix} X & 0 \\ 0 & \alpha Y \end{bmatrix}$. Then $\bar{Z} = \bar{Z}^T > 0$ as $X > 0$ and $Y > 0$. Using Lemma B.2, we can derive

$$\bar{A}^T\bar{Z} + \bar{Z}\bar{A} + \bar{A}_p^T\bar{Z}\bar{A}_p + \bar{Q} < 0, \tag{B.38}$$

which implies

$$\begin{bmatrix} A+BK & -BK \\ 0 & A+GC \end{bmatrix}^T \begin{bmatrix} X & 0 \\ 0 & \alpha Y \end{bmatrix} + \begin{bmatrix} X & 0 \\ 0 & \alpha Y \end{bmatrix} \begin{bmatrix} A+BK & -BK \\ 0 & A+GC \end{bmatrix} \tag{B.39}$$

$$+ \begin{bmatrix} A_p & 0 \\ A_p & 0 \end{bmatrix}^T \begin{bmatrix} X & 0 \\ 0 & \alpha Y \end{bmatrix} \begin{bmatrix} A_p & 0 \\ A_p & 0 \end{bmatrix} + \begin{bmatrix} Q + K^T RK & -K^T RK \\ -K^T RK & K^T RK \end{bmatrix} < 0. \tag{B.40}$$

Equivalently,

$$\begin{bmatrix} \Lambda_1 + \alpha A_p^T Y A_p & -X(BK) - K^T RK \\ -(BK)^T X - K^T RK & \alpha\Lambda_2 + K^T RK \end{bmatrix} < 0, \tag{B.41}$$

where

$$\Lambda_1 = (A+BK)^T X + X(A+BK) + A_p^T X A_p + Q + K^T RK, \tag{B.42}$$

$$\Lambda_2 = (A+GC)^T Y + Y(A+GC). \tag{B.43}$$

Hence, the theorem is proved. □

Theorem B.2. *[Gao and Shi (2013)] For the closed-loop system* (B.30)*, if there exist matrices*
(i) $U = U^T > 0$*,* Z *and a positive scalars* ρ *such that:*

$$\begin{bmatrix} \Gamma_{UV} & UA_p^T & U & Z^T \\ A_p U & -U & 0 & 0 \\ U & 0 & -Q^{-1} & 0 \\ Z & 0 & 0 & -R^{-1} \end{bmatrix} < 0, \tag{B.44}$$

where

$$\Gamma_{UV} := UA^T + AU + Z^T B^T + BZ + \rho I_n,$$

(ii) $Y = Y^T > 0$*,* W *and a positive scalar* β *such that:*

$$\begin{bmatrix} \Gamma_{XY} & -\beta X(BK) - \beta K^T RK \\ -\beta (BK)^T X - \beta K^T RK & A^T Y + YA + WC + C^T W^T + \beta K^T RK \end{bmatrix} < 0, \tag{B.45}$$

where

$$\Gamma_{XY} := -\beta\rho X^T X + A_p^T Y A_p,$$

then control gain K and observer gain G can be obtained as follows:

$$K = ZU^{-1}, \tag{B.46}$$

$$G = Y^{-1} W. \tag{B.47}$$

Proof. **Proof (i):** Compared to the Theorem 3 of [Gao and Shi (2013)],

$$\Lambda_1 < -\rho X^T X, \tag{B.48}$$

where

$$\Lambda_1 = (A+BK)^T X + X(A+BK) + A_p^T X A_p + Q + K^T RK. \tag{B.49}$$

Multiplying X^{-1} on the left-hand side and on the right-hand side of (B.48), and letting $U = X^{-1}$, we can get

$$U(A+BK)^T + (A+BK)U + UA_p^T U^{-1} + A_p U + \rho I_n + U(Q+K^T RK)U < 0. \qquad \text{(B.50)}$$

Now letting $Z = KU$ and using Schur complement inequality, (B.50) is implied by

$$\begin{bmatrix} \Gamma_{UV} & UA_p^T & U & Z^T \\ A_p U & -U & 0 & 0 \\ U & 0 & -Q^{-1} & 0 \\ Z & 0 & 0 & -R^{-1} \end{bmatrix} < 0, \qquad \text{(B.51)}$$

where

$$\Gamma_{UV} := UA^T + AU + Z^T B^T + BZ + \rho I_n.$$

Hence inequality (i) is proved.
Proof (ii): Applying $\Lambda_1 < -\rho X^T X$ from (B.48) into (B.35), we can write

$$\begin{bmatrix} -\rho X^T X + \alpha A_p^T Y A_p & -X(BK) - K^T RK \\ -(BK)^T X - K^T RK & \alpha \Lambda_2 + K^T RK \end{bmatrix} < 0, \qquad \text{(B.52)}$$

where

$$\Lambda_2 = (A+GC)^T Y + Y(A+GC).$$

Let $W = Y^{-1}G$, and $\alpha = 1/\beta$, we can obtain

$$\begin{bmatrix} \Gamma_{XY} & -\beta X(BK) - \beta K^T RK \\ -\beta (BK)^T X - \beta K^T RK & A^T Y + YA + WC + C^T W^T + \beta K^T RK \end{bmatrix} < 0, \qquad \text{(B.53)}$$

where

$$\Gamma_{XY} := -\beta \rho X^T X + A_p^T Y A_p.$$

This competes the proof. □

Furthermore, the corresponding value of the cost function (B.24) satisfies the following inequality:

$$J < \mathbb{E}[x_0^T U^{-1} x_0]. \qquad \text{(B.54)}$$

Consequently, solving the following optimization problem allows us to determine the optimal cost bound.

$$J^* = \min \psi, \qquad \text{(B.55)}$$

subject to

$$\begin{bmatrix} -\psi & x_0^T \\ x_0 & -U \end{bmatrix} < 0. \qquad \text{(B.56)}$$

B.2.3 Future works

In this section, I would like to formulate a problem of stochastic incentive Stackelberg game through observer-based control design. However, the results and discussions on this issue will be future investigations. Consider a stochastic system governed by the Itô differential equation defined by

$$dx(t) = \big[Ax(t) + B_0u_0(t) + B_1u_1(t)\big]dt + A_px(t)dw(t),\ \ x(0) = x^0, \tag{B.57a}$$

$$y(t) = Cx(t) \tag{B.57b}$$

where $u_0(t) \in \mathbb{R}^{m_0}$ denotes the leader's control input. $u_1(t) \in \mathbb{R}^{m_1}$ denotes the follower's control input. The definitions of the other variables are the same as those in stochastic system (B.23). The coefficient matrices A, B_0, B_1, A_p and C are of suitable dimensions. In the following, we use P_0 to represent the leader and P_1 to represent the follower.

On the other hand, the cost functions of P_0 and P_1 are correspondingly given by

$$J_0(u_0,\ u_1) = \frac{1}{2}\mathbb{E}\left[\int_0^\infty \Big\{x^T(t)Q_0x(t) + u_0^T(t)R_{00}u_0(t) + u_1^T(t)R_{01}u_1(t)\Big\}dt\right], \tag{B.58a}$$

$$J_1(u_0,\ u_1) = \frac{1}{2}\mathbb{E}\left[\int_0^\infty \Big\{x^T(t)Q_1x(t) + u_0^T(t)R_{10}u_0(t) + u_1^T(t)R_{11}u_1(t)\Big\}dt\right], \tag{B.58b}$$

where $Q_1 = Q_1^T \geq 0$, $R_{00} = R_{00}^T > 0$, $R_{01} = R_{01}^T \geq 0$, $R_{10} = R_{10}^T \geq 0$ and $R_{11} = R_{11}^T > 0$, are of suitable dimensions. The game process to determine the incentive Stackelberg strategy set is as follows.

(i) In the stochastic system (B.57), the leader P_0 achieves an observer-based output feedback strategy of team-optimal condition (B.59),

$$J_0(u_0^*,\ u_1^*) = \min_{u_0,\ u_1} J_0(u_0,\ u_1). \tag{B.59}$$

(ii) The leader announces the strategy to the follower in advance through the following estimated state observer pattern:

$$u_0(t) = u_0(t, \hat{x}(t), u_1) = \Lambda\hat{x}(t) + \Xi u_1(t), \tag{B.60}$$

where $\Lambda \in \mathbb{R}^{m_0 \times n}$ and $\Xi \in \mathbb{R}^{m_0 \times m_1}$ are strategy parameter matrices and

$$d\hat{x}(t) = [A\hat{x}(t) + B_cu_c(t)]dt + G[\hat{y}(t) - y(t)]dt, \quad \hat{x}(0) = x_0, \tag{B.61a}$$

$$\hat{y}(t) = C\hat{x}(t), \tag{B.61b}$$

$$u_c(t) = \begin{bmatrix} K_{c0} \\ K_{c1} \end{bmatrix} \hat{x}(t) = K_c\hat{x}(t), \tag{B.61c}$$

with

$$B_c := \begin{bmatrix} B_0 & B_1 \end{bmatrix}, \quad u_c := \mathbf{col} \begin{bmatrix} u_0 & u_1 \end{bmatrix}.$$

Furthermore, the cost functional in this case is

$$J_0 \left(u_c(t) \right) := \frac{1}{2} \mathbb{E} \left[\int_0^{\infty} \left\{ \hat{x}^T(t) Q_0 \hat{x}(t) + u_c^T(t) R_c u_c(t) \right\} dt \right], \tag{B.62}$$

where

$$R_c := \mathbf{block\ diag} \begin{bmatrix} R_{00} & R_{01} \end{bmatrix}.$$

(iii) The follower determines the optimal strategy $u_1^*(t)$ through the output feedback by means of the state observer responding to the announced strategy of the leader. i.e., to find $u_1^*(t)$ with

$$d\hat{x}(t) = [A\hat{x}(t) + B_0[\Lambda\hat{x}(t) + \Xi u_1(t)] + B_1 u_1(t)]dt + G[\hat{y}(t) - y(t)]dt, \quad \hat{x}(0) = x_0, \tag{B.63a}$$

$$\hat{y}(t) = C\hat{x}(t), \tag{B.63b}$$

$$u_1(t) = K_1 \hat{x}(t). \tag{B.63c}$$

The cost functional in this case is

$$J_1(u_1) = \frac{1}{2} \mathbb{E} \left[\int_0^{\infty} \Big\{ \hat{x}^T(t) Q_1 \hat{x}(t) + [\Lambda\hat{x}(t) + \Xi u_1(t)]^T R_{10} [\Lambda\hat{x}(t) + \Xi u_1(t)] + u_1^T(t) R_{11} u_1(t) \Big\} dt \right]. \tag{B.64}$$

(iv) Using the equivalence relation $K_{c1} = K_1$, the leader determines unknown parameters and the incentive Stackelberg strategy

$$u_0^*(t) = u_0^*(t, \hat{x}(t), u_1) = \Lambda\hat{x}(t) + \Xi u_1^*(t), \tag{B.65}$$

to achieve the team optimal solution $(u_0^*,\ u_1^*)$.

Remark B.3. *As we saw the state feedback control, when we consider incentive Stackelberg games, the coefficients B_0 and B_1 associated with the control inputs should be square matrices. The incentive parameter Ξ was determined by the following matrix algebraic equation:*

$$\Xi^T (B_0^T P_1 - R_{10} R_{01}^{-1} B_0^T P_c) = R_{11} R_{01}^{-1} B_1^T P_c - B_1^T P_1, \tag{B.66}$$

where $P_1, P_c \in \mathbb{R}^{n\times n}$ are symmetric positive semi-definite matrices. The incentive parameter Ξ can be uniquely determined if and only if $(B_0^T P_1 - R_{10}R_{01}^{-1}B_0^T P_c)$ is non-singular. If B_0 is not a square matrix, the statement does not hold. We can also observe that R_{01} should have the same size as B_0. On the right side, the same dimension of R_{01} and P_c means that the dimensions of B_1 must be the same and square. However, with the observer-based output feedback strategy, it can be seen from the preliminary results that no relationship like (B.66) *will occur. Therefore, to avoid this serious drawback, an observer-based output feedback design can be considered.*

B.3 H_∞ constrained nonlinear stochastic LQ control

B.3.1 Preliminaries

Consider a nonlinear stochastic system governed by the Itô differential equation defined by

$$dx(t) = \left[f(x) + \sum_{j=1}^{N} g_j(x)u_j(t) + h(x)v(t)\right]dt + r(x)dw, \;\; x(0) = x^0, \tag{B.67a}$$

$$z(t) = \mathbf{col}\left[\; Cx(t) \quad u_1(t) \quad \cdots \quad u_N(t) \;\right], \tag{B.67b}$$

where $f(0) = 0$ and $h(0) = 0$. Here, $x(t) \in \mathbb{R}^n$ represents the state vector. $z(t) \in \mathbb{R}^{n_z}$ represents the controlled output. $u_i(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\, \mathbb{R}^{m_i})$, $i = 1,\, \dots\, ,N$ represent the i-th control inputs. $v(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\, \mathbb{R}^{m_v})$ represent the exogenous disturbance signal. $w(t) \in \mathbb{R}$ represents a one-dimensional standard Wiener process defined in the filtered probability space $(\Omega,\, \mathscr{F},\, \mathscr{P},\, \mathscr{F}_t)$ with $\mathscr{F}_t = \sigma\{w(s) \,:\, 0 \le s \le t\}$ [Zhang et al. (2006), Zhang et al. (2008)].

Now, the finite horizon nonlinear stochastic H_∞ control with multiple decision makers that is based on a stochastic Nash game is given below.

Definition B.5. *For any given $\gamma > 0$, $0 < T_f < \infty$, $v(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\, \mathbb{R}^{m_v})$, find, if possible, a state feedback strategy set $u_i(t) = u_i^*(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\, \mathbb{R}^{m_i})$, $i = 1,\, \dots\, ,N$, such that*
(i) The trajectory of the closed-loop system (B.67) *satisfies*

$$0 \le J_v(u_1^*,\, \dots\, ,u_N^*,\, v^*,\, x^0) \le J_v(u_1^*,\, \dots\, ,u_N^*,\, v,\, x^0), \tag{B.68}$$

where

$$J_v(u_1,\, \dots\, ,u_N,\, v,\, x^0) = \mathbb{E}\left[\int_0^{T_f} [\gamma^2\|v(t)\|^2 - \|z(t)\|^2]dt\right], \tag{B.69}$$

$$\|z(t)\|^2 = x^T(t)C^T Cx(t) + \sum_{j=1}^{N} u_j^T(t)u_j(t) \tag{B.70}$$

for $\forall v(t) \neq 0 \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\ \mathbb{R}^{m_v})$.
(ii) When the worst case disturbance $v^*(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\ \mathbb{R}^{m_v})$*, if existing, is implemented in (1),* $u_i(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f],\ \mathbb{R}^{m_i})$*,* $i = 1,\ \dots\ ,N$ *satisfy the following the Nash equilibria defined by*

$$J_i(u_1^*,\ \dots\ ,u_N^*,\ v^*,\ x^0) \leq J_i(u_1^*,\ \dots\ ,\ u_{i-1}^*,\ u_i,\ u_{i+1}^*,\ \dots\ ,u_N^*,\ v^*,\ x^0), \tag{B.71}$$

where

$$J_i(u_1,\ \dots\ ,u_N,\ v,\ x^0) = \mathbb{E}\left[\int_0^{T_f} [x^T(t)Q_i x(t) + u_i^T(t)R_i u_i(t)]dt\right] \tag{B.72}$$

with $Q_i = Q_i^T \geq 0$ *and* $R_i = R_i^T > 0$.

That is, the considered H_∞ control problem for stochastic nonlinear systems with multiple decision makers is to find the v^* and u_i^* such that the inequalities (B.68) and (B.71) hold, respectively.

Theorem B.3. *[Mukaidani et al. (2015b)] Consider the following cross-coupled Hamilton-Jacobi-Bellman (HJB) equations:*

$$\frac{\partial V_v^T}{\partial t} + \frac{\partial V_v^T}{\partial x}\hat{f}_K(x) - \frac{\gamma^{-2}}{4}\cdot\frac{\partial V_v^T}{\partial x}h(x)h^T(x)\frac{\partial V_v}{\partial x} - \hat{m}^T(t,x)\hat{m}(t,x) + \frac{1}{2}r^T(x)\frac{\partial^2 V_v}{\partial x^2}r(x) = 0, \tag{B.73a}$$

$$\frac{\partial V_i^T}{\partial t} + \frac{\partial V_i^T}{\partial x}\hat{f}_{-i}(x) + x^T(t)Q_i x(t) - \frac{1}{4}\cdot\frac{\partial V_i^T}{\partial x}g_i(x)R_i^{-1}g_i^T(x)\frac{\partial V_i}{\partial x} + \frac{1}{2}r^T(x)\frac{\partial^2 V_i}{\partial x^2}r(x) = 0, \tag{B.73b}$$

where $i = 1,\ \dots\ ,N,\ V_v = V_v(t,x),\ V_i = V_i(t,x),\ V_v(T_f,x(T_f)) = 0,\ V_i(T_f,x(T_f)) = 0,$

$$\hat{f}_K(x(t)) := f(x(t)) + \sum_{j=1}^{N} g_j(x)K_j(t,x),$$

$$\hat{f}_{-i}(x(t)) := f(x(t)) + \sum_{j=1,\ i\neq i}^{N} g_j(x)K_j(t,x) + h(x(t))K_v(t,x),$$

$$\hat{m}^T\hat{m} := x^T(t)C^TCx(t) + \sum_{j=1}^{N} K_j^T(t,x)K_j(t,x).$$

Suppose there exist a set of solutions $(V_v,\ V_1,\ \dots\ ,V_N)$ *then the finite horizon Nash based* H_∞ *control has a strategy set*

$$v^*(t,x) = K_v^*(t,x) = -\frac{\gamma^{-2}}{2}h^T\frac{\partial V_v^*(t,x)}{\partial x}, \tag{B.74a}$$

$$u_i^*(t,x) = K_i^*(t,x) = -\frac{1}{2}R_i^{-1}g_i^T\frac{\partial V_i^*(t,x)}{\partial x}. \tag{B.74b}$$

Proof. Let us consider the following H_∞ control problem for the closed-loop stochastic system with arbitrary strategies $u_i(t,x) = K_i(t,x)$, $i = 1, \ldots, N$.

$$dx(t) = \left[\hat{f}_K(x(t)) + h(x(t))v(t)\right] dt + r(x(t))dw(t), \tag{B.75a}$$

$$z(t) = \hat{m} := \left[\begin{array}{cccc} x^T(t)C^T & K_1^T(t,x) & \cdots & K_N^T(t,x) \end{array}\right]^T. \tag{B.75b}$$

Applying Lemma 2.1 of [Zhang et al. (2006)] to the stochastic system (B.75) yields the HJB equation (B.73a) and a worst case disturbance is (B.74a). Furthermore, in order to apply LQ problem that is defined by [Zhang et al. (2006)], the following regulator problem is considered.

$$\begin{aligned} &\text{minimize } J_{-i}(u_i), \\ &\text{s.t. } dx(t) = [\hat{f}_{-i}(x(t)) + g_i(x(t))u_i(t)]dt + r(x(t))dw(t), \end{aligned} \tag{B.76}$$

where

$$\begin{aligned} J_{-i}(u_i) := \quad & J_i(K_1^* \ldots, K_{i-1}^*, u_i, K_{i+1}^*, \ldots, K_N^*, v^*, x^0) \\ = \quad & \mathbb{E}\left[\int_0^{T_f} \left[x^T(t)Q_i x(t) + u_i^T(t,x)R_i u_i(t,x)\right] dt\right]. \end{aligned}$$

If there exists a nonnegative Lyapunov function $V_i(t,x)$ solving the HJB equation (B.73b). Thus, we have

$$J_{-i}(K_i^*) \geq \mathbb{E}[V_i(0,x^0)] \tag{B.77}$$

with the optimal strategy (B.74b) can be derived, respectively. □

B.3.2 Future works

In this section, I would like to formulate a problem of H_∞-constrained incentive Stackelberg game with stochastic nonlinear system dynamics. However, the results and discussions on this issue will be future investigations. Consider a nonlinear stochastic system governed by the Itô differential equation defined by

$$dx(t) = \left[f(x) + \sum_{j=1}^{N}\left[g_{0j}(x)u_{0j}(t) + g_j(x)u_j(t)\right] + h(x)v(t)\right]dt + r(x)dw, \ \ x(0) = x^0, \tag{B.78a}$$

$$z(t) = \mathbf{col}\left[\begin{array}{ccccc} Cx(t) & \boldsymbol{u}_0(t) & u_1(t) & \cdots & u_N(t) \end{array}\right], \tag{B.78b}$$

$$\boldsymbol{u}_0(t) = \mathbf{col}\left[\begin{array}{ccc} u_{01}(t) & \cdots & u_{0N}(t) \end{array}\right], \tag{B.78c}$$

where $f(0) = 0$ and $h(0) = 0$. Here, $\boldsymbol{u}_0(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f], \mathbb{R}^{m_0})$, $m_0 = \sum_{j=1}^N m_{0j}$ with $u_{0j}(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f], \mathbb{R}^{m_{0j}})$ denotes the leader's control input. $u_i(t) \in \mathscr{L}^2_{\mathscr{F}}([0,T_f], \mathbb{R}^{m_i})$, $i = 1, \ldots, N$

denotes the ith follower's control input. In the following, we use P_0 to represent the leader and P_i, $i = 1, \ldots, N$ to represent the ith follower. The definitions of the other variables are the same as those in stochastic system (B.67).

On the other hand, the cost functions of P_0 and P_i, $i = 1, 2, \ldots, N$ are given by

$$J_0(u_{01}, \ldots, u_{0N}, u_1, \ldots, u_N, v) = \frac{1}{2}\mathbb{E}\left[\int_0^{T_f}\left\{x^T(t)Q_0x(t) + \sum_{i=1}^N\left[u_{0i}^T(t)R_{00i}u_{0i}(t) + u_i^T(t)R_{0i}u_i(t)\right]\right\}dt\right], \quad \text{(B.79a)}$$

$$J_i(u_{01}, \ldots, u_{0N}, u_1, \ldots, u_N, v) = \frac{1}{2}\mathbb{E}\left[\int_0^{T_f}\left\{x^T(t)Q_ix(t) + u_{0i}^T(t)R_{0ii}u_{0i}(t) + u_i^T(t)R_{ii}u_i(t)\right\}dt\right], \ i = 1, \ldots, N, \quad \text{(B.79b)}$$

where $Q_0 := C^TC$, $Q_i = Q_i^T \geq 0$, $R_{00i} = R_{00i}^T > 0$, $R_{0i} = R_{0i}^T \geq 0$, $R_{0ii} = R_{0ii}^T \geq 0$ and $R_{ii} = R_{ii}^T > 0$, $i = 1, \ldots, N$ are of suitable dimensions, and the entries are piece-wise continuous functions of time on the fixed interval $[0, T_f]$.

The game process to determine the incentive Stackelberg strategy set is as follows:
(i) The leader announces a strategy ahead of time to the followers with the following feedback pattern.

$$u_{0i}(t) = u_{0i}(t, x(t), u_i) = \Gamma_ix(t) + \Xi_iu_i(t), \quad \text{(B.80)}$$

for $i = 1, \ldots, N$, where $\Gamma_i \in \mathbb{R}^{m_{0i}\times n}$ and $\Xi_i \in \mathbb{R}^{m_{0i}\times m_i}$ are strategy parameter matrices of suitable dimensions. Moreover, their components are piece-wise continuous functions of time on the interval $[0, T_f]$.
(ii) The followers determine their strategies to achieve a Nash equilibrium by responding to the announced strategy of the leader.
(iii) The leader determines the incentive Stackelberg strategy

$$u_{0i}^*(t) = \Gamma_ix(t) + \Xi_iu_i^*(t), \quad \text{(B.81)}$$

for $i = 1, \ldots, N$ to achieve the team optimal solution $(\boldsymbol{u}_0^*, u_1^*, \ldots, u_N^*)$, which is associated with the Nash equilibrium strategy $u_i^*(t)$ for $i = 1, \ldots, N$ of the followers.

The finite-horizon H_∞-constrained incentive Stackelberg game with multiple non-cooperative followers can be formulated as follows.

For any disturbance attenuation level $\gamma > 0$, $0 < T_f < \infty$, we need to find an incentive strategy of P_0 by (B.81) and a closed-loop Nash strategy of P_i by

$$u_i^*(t) := K_i(t), \quad i = 1, \ldots, N,$$

considering the worst-case disturbance $v^*(t) := F_\gamma(t)$ such that

(i) The trajectory of the closed-loop system (B.78) satisfies the following team-optimal condition (B.82a) along with H_∞ constraint condition (B.82b),

$$J_0(x_0, \boldsymbol{u}_0^*, u_1^*, \dots, u_N^*, v^*) \le J_0(x_0, \boldsymbol{u}_0, u_1, \dots, u_N, v^*), \tag{B.82a}$$

$$0 \le J_v(x_0, \boldsymbol{u}_0^*, u_1^*, \dots, u_N^*, v^*) \le J_v(x_0, \boldsymbol{u}_0^*, u_1^*, \dots, u_N^*, v). \tag{B.82b}$$

(ii) A set of decision $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i = 1,\ \dots, N$ satisfying the following Nash equilibrium inequality:

$$J_i(x_0, \boldsymbol{u}_0^*, u_1^*, \dots, u_N^*, v^*) \le J_i(x_0, \lambda_{-i}^*(u_{0i}), \lambda_{-i}^*(u_i), v^*) \tag{B.83}$$

with $\lambda_{-i}^*(\alpha) := (\lambda_1^*, \dots, \lambda_{i-1}^*,\ \alpha,\ \lambda_{i+1}^*, \dots, \lambda_N^*)$.

Then, the strategy-set $(u_{0i}^*,\ u_i^*) \in \mathbb{R}^{m_{0i}+m_i},\ i = 1,\ \dots, N$ constitutes both a team-optimal incentive Stackelberg strategy with the H_∞ constraint of the leader and Nash equilibrium strategies of the followers for a two-level hierarchical game.

Remark B.4. *It should be noted that to implement the proposed method for the nonlinear stochastic systems, HJB equations like* (B.73) *will appear, and it is more difficult to find appropriate solutions numerically than that of the linear systems.*

Appendix C

Stochastic H_∞-Control for Small γ

If the disturbance attenuation level is too small, the H_∞-control design described in our thesis will be invalid. To solve such problems [Pan and Başar (1993)] characterizes a class of stabilizing controllers. Furthermore, the minimum value of $\gamma = \gamma_{\min}$ can be found until the system is stable. Consider the following stochastic linear system:

$$dx(t) = [Ax(t) + Bu(t) + Ev(t)]\,dt + A_p x(t)dw, \qquad x(0) = x^0 \tag{C.1a}$$

$$z(t) = \mathbf{col}\left[Cx(t) \quad Du(t)\right], \quad D^T D = I, \tag{C.1b}$$

where, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})$ is the control input, $v(t) \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v})$ is the deterministic disturbance, $w(t) \in \mathbb{R}$ is a one-dimensional wiener process and $z(t) \in \mathbb{R}^{n_z}$ is the controlled output.

With this system, we associate the standard quadratic performance index:

$$J_u(u,\ v) = \|z\|_2^2 = \mathbb{E}\left[\int_0^\infty \left\{x^T C^T Cx + u^T u\right\} dt\right]. \tag{C.2}$$

The H_∞-optimal control problem is the minimization of the quantity

$$\|\mathscr{L}\|_\infty = \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\|z\|_2}{\|v\|_2}. \tag{C.3}$$

The derivation of a feedback controller $u = Kx$ will ensure a performance within a given neighborhood of the infimum of (C.3). Let this minimum value be represented by $\gamma_{\min}$, i.e.

$$\gamma_{\min} = \inf_{u \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})} \ \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\|z\|_2}{\|v\|_2}. \tag{C.4}$$

We can associate this to the linear-quadratic differential game [Başar and Bernhard (2008)] with worst-case disturbance problem, which has the cost functional

$$J(u,\ v) = \mathbb{E}\left[\int_0^\infty (\|z\|^2 - \gamma^2 \|v\|^2) dt\right]. \tag{C.5}$$

Definition C.1. *[Başar and Bernhard (2008)] A strategy pair* $(u^*,\ v^*) \in \Gamma_u \times \Gamma_v$ *is in saddle-point equilibrium if*

$$J(u^*,\ v) \leq J(u^*,\ v^*) \leq J(u,\ v^*)$$

for all $(u^*, v) \in \Gamma_u \times \Gamma_v$ *and* $(u,\ v^*) \in \Gamma_u \times \Gamma_v$*, where* $\Gamma_u \times \Gamma_v$ *means a product vector space.*

For each $\gamma > \gamma_{\min}$, the differential game allows the saddle-point controller design [Başar and Bernhard (2008)] with state feedback law. The design can also be transferred as a state feedback Nash equilibrium strategy [Zhang and Chen (2004)] by changing the sign of the disturbance performance (C.5) as

$$J_v(u,v) = \mathbb{E}\left[\int_0^\infty (\gamma^2\|v\|^2 - \|z\|^2)dt\right]. \tag{C.6}$$

Due to the numerical stiffness, the computation of $\gamma_{\min}$ and the corresponding H_∞-optimal controller for small values of $\gamma > 0$ have serious difficulties.

The following theorem can be used to design the controller whether such problems of small disturbance attenuation levels occur.

Theorem C.1. *For the system* (C.1)*, suppose the generalized algebraic Riccati equation (GARE):*

$$A^T Z + ZA + A_p^T Z A_p - ZSZ + C^T C = 0, \tag{C.7}$$

with

$$S := BB^T - \gamma^{-2} EE^T,$$

has the solution $Z > 0 > 0$*. If* $[A,\ A_p|C]$ *is exactly observable, then the stochastic* H_2/H_∞ *control problem admits a solution set:*

$$u^*(t) = K^* x(t) = -B^T Z x(t), \tag{C.8}$$

$$v^*(t) = K_\gamma^* x(t) = \gamma^{-2} E^T Z x(t). \tag{C.9}$$

Proof. Let $u(t) = u^*(t) = -B^T Z x(t)$, then by Lemma 5 of [Zhang and Chen (2004)], there exist a solution $Z > 0$ to the following ARE:

$$Z(A - BB^T Z) + (A - BB^T Z)^T Z + A_p^T Z A_p + ZBB^T Z + \gamma^{-2} ZEE^T Z + C^T C = 0, \tag{C.10}$$

which is the same as GARE (C.7).

Now we have to show

(i) $\|\mathscr{L}\|_\infty < \gamma$.

Substituting $u(t) = u^*(t) = -B^T Zx(t)$ into (C.1) gives

$$\begin{cases} dx(t) = \left\{(A - BB^T Z)x(t) + Ev(t)\right\} dt + A_p x(t) dw(t), \\ z(t) = \begin{bmatrix} Cx(t) \\ -DB^T Zx(t) \end{bmatrix}, \end{cases} \tag{C.11}$$

where $x(0) = x_0$. Applying Ito's formula to (C.11) and considering GARE (C.10), we have

$$\mathbb{E}\left[\int_0^\infty d(x^T Zx)\right] = \mathbb{E}\left[\int_0^\infty \left\{\left((A - BB^T Z)x + Ev\right)(Zx + x^T Z) + x^T A_p^T Z A_p x\right\} dt\right]$$

$$\begin{aligned}
\text{or, } -x_0^T Z x_0 = \mathbb{E}\Big[\int_0^\infty \Big\{ & x^T (Z(A - BB^T Z) + (A - BB^T Z)^T Z + A_p^T Z A_p) x \\
& + v^T E^T Z x + x^T Z E v \Big\} dt\Big] \\
= \mathbb{E}&\left[\int_0^\infty \left\{x^T \tilde{A}_2^T \tilde{A}_2 x + v^T E^T Zx + x^T ZEv\right\} dt\right] \\
= \mathbb{E}&\left[\int_0^\infty \left\{x^T (C^T C + \gamma^{-2} ZEE^T Z + ZBB^T Z)x + v^T E^T Zx + x^T ZEv\right\} dt\right] \\
= \mathbb{E}&\left[\int_0^\infty \left\{z^T z + \gamma^2 {v^*}^T v^* + v^T E^T Zx + x^T ZEv\right\} dt\right] \\
& [\text{suppose, } v^*(t) = \gamma^{-2} E^T Zx(t)]
\end{aligned}$$

$$\begin{aligned}
\text{or, } \mathbb{E}\left[\int_0^\infty \left\{\gamma^2 v^T v - z^T z\right\} dt\right] &= x_0^T Z x_0 + \mathbb{E}\left[\int_0^\infty \left\{\gamma^2 v^T v + \gamma^2 {v^*}^T v^* - \gamma^2 v^T v^* - \gamma^2 {v^*}^T v\right\} dt\right] \\
&= x_0^T Z x_0 + \gamma^2 \mathbb{E}\left[\int_0^\infty (v - v^*)^T (v - v^*) dt\right]
\end{aligned} \tag{C.12}$$

So

$$\begin{aligned}
J_v(u^*,\ v) &= \mathbb{E}\left[\int_0^\infty \left\{\gamma^2 v^T v - z^T z\right\} dt\right] \\
&= x_0^T Z x_0 + \gamma^2 \mathbb{E}\left[\int_0^\infty (v - v^*)^T (v - v^*) dt\right] \geq J_v(u^*,\ v^*) = x_0^T Z x_0.
\end{aligned} \tag{C.13}$$

Now, if we define an operator $\mathscr{L}_1 v = v - v^*$, then form (C.13) we have (for $x(0) = x_0 = 0$):

$$J_v(u^*,\ v) = \gamma \|v\|^2 - \|z\|^2 = \gamma^2 \|\mathscr{L}_1 v\|^2 \geq \varepsilon \|v\|^2 > 0$$

, for some $\varepsilon > 0$, which yields $\|\mathscr{L}\|_\infty < \gamma$.

(ii) u^* minimizes the output energy $\|z\|_2^2$ when v^* applied in (C.1), i.e.,

$$u^* = \arg\min_u J_u(u,\ v^*), \qquad \forall u \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n}_{u}).$$

When worst-case disturbance $v = v^*(t) = \gamma^{-2} E^T Zx(t)$ is applied to (C.1), we have

$$\begin{cases} dx(t) = \left\{(A + \gamma^{-2} EE^T Z)x(t) + Bu(t)\right\} dt + A_p x(t) dw(t) \\ z(t) = \begin{bmatrix} Cx(t) \\ Du(t) \end{bmatrix}, \end{cases} \tag{C.14}$$

where $x(0) = x_0$. Now the H_2 optimization problem becomes a standard stochastic LQ optimal control problem, so we can write [Rami and Zhou (2000)]

On the other hand, let $v(t) = v^*(t) = K_\gamma^* x(t) = \gamma^{-2} E^T Z x(t)$, then by Lemma 6 of [Zhang and Chen (2004)], there exist a solution $Z > 0$ to the following ARE:

$$Z(A+\gamma^{-2}ZEE^TZ) + (A+\gamma^{-2}ZEE^TZ)^TZ + A_p^TZA_p + ZBB^TZ - \gamma^{-2}ZEE^TZ + C^TC = 0, \tag{C.15}$$

which is the same as GARE (C.7). Applying Ito's formula in (C.14) considering (C.15) we get,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\infty d(x^TZx)\right] = \mathbb{E}\left[\int_0^\infty \left\{\left((A+\gamma^{-2}EE^TZ)x + Bu\right)(Zx + x^TZ) + x^TA_p^TZA_px\right\} dt\right] \\
&\text{or, } -x_0^TZx_0 = \mathbb{E}\left[\int_0^\infty \left\{x^TZBB^TZx - x^TC^TCx + u^TB^TZx + x^TZBu\right\} dt\right] \text{ [by (C.15)]} \\
&\text{or, } \mathbb{E}\left[\int_0^\infty \left\{x^TC^TCx + u^Tu\right\} dt\right] \\
&\qquad = x_0^TZx_0 + \mathbb{E}\left[\int_0^\infty \left\{u^Tu + u^{*^T}u^* + u^Tu^* - u^Tu^* - u^{*^T}u\right\} dt\right] \\
&\text{or, } J_u(u,\ v^*) = x_0^TZx_0 + \mathbb{E}\left[\int_0^\infty (u-u^*)^T(u-u^*)dt\right]. \qquad \text{(C.16)}
\end{aligned}$$

If we put $u = u^*$, then form (C.16) we get

$$J_u(u,\ v^*) \geq J_u(u^*,\ v^*) = x_0^TZx_0. \tag{C.17}$$

It can be observed from both (C.10) and (C.15) that these equations can be simplified as the GARE (C.7). So, the solution of GARE (C.7) can be written as $Z > 0$. □

Using above results we can find from (C.4)

$$\begin{aligned}
\gamma_{\min} &= \inf_{u \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})} \ \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\sqrt{\mathbb{E}\left[\int_0^\infty \{x^TC^TCx + u^Tu\}\, dt\right]}}{\sqrt{\mathbb{E}\left[\int_0^\infty v^Tv\, dt\right]}} \\
&= \inf_{u \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_u})} \ \sup_{\substack{v \in \mathscr{L}^2_{\mathscr{F}}(\mathbb{R}_+,\ \mathbb{R}^{n_v}) \\ v \neq 0,\ x^0 = 0}} \frac{\sqrt{\mathbb{E}\left[\int_0^\infty \{x^T[C^TC + K^{*T}K^*]x\}\, dt\right]}}{\sqrt{\mathbb{E}\left[\int_0^\infty x^TK_\gamma^TK_\gamma x\, dt\right]}}.
\end{aligned}$$

Numerical example

In order to demonstrate the efficiency of our proposed strategies, a numerical example is investigated. Let us consider the following system matrices of the system (C.1):

$$A = \begin{bmatrix} -0.52 & 1.12 & 0 \\ 0 & -0.24 & 1 \\ 0.23 & 0.85 & -0.16 \end{bmatrix}, \; A_p = 0.1A, \; x(0) = \begin{bmatrix} 1 \\ 0.5 \\ -0.6 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.15 \\ 0.12 \\ 3.55 \end{bmatrix}, \; E = \begin{bmatrix} 0.23 \\ -0.52 \\ 0.28 \end{bmatrix}, \; C = [1 \quad 1 \quad 1], \; \gamma = 3.$$

If we apply our proposed method to the system (C.1), we get the following numerical results.

$$Z = \begin{bmatrix} 2.3417e-01 & 2.8373e-01 & 2.4301e-01 \\ 2.8373e-01 & 4.3796e-01 & 3.7839e-01 \\ 2.4301e-01 & 3.7839e-01 & 3.3864e-01 \end{bmatrix},$$

$$K = \begin{bmatrix} -9.3184e-01 & -1.4384e+00 & -1.2840e+00 \end{bmatrix},$$

$$K_\gamma = \begin{bmatrix} -2.8487e-03 & -6.2811e-03 & -5.1173e-03 \end{bmatrix}.$$

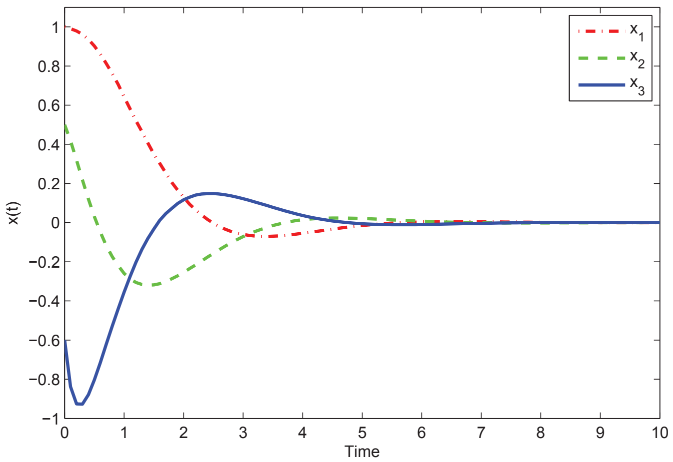


Fig. C.1: Trajectory of the state.

The MATLAB code is developed on the basis of Lyapunov iterations which converges to the required solutions of GARE (C.7) with an accuracy of 1.0e − 14 order only after 7 iterations. It can be observed form the response depicted in Fig. C.1 that the state attains the mean-square stable.

γ	3	1.5	$\cdots$	0.09375	0.140625	$\cdots$	0.1532178
Solution	exists	exist	$\cdots$	does not exist	does not exist	$\cdots$	exists ($\gamma_{\min}$)

Table C.1: Bisection method for finding $\gamma_{\min}$.

By applying the bisection method, it can be found that the minimum value of $\gamma = \gamma_{\min} = 0.1532178$ with an accuracy of $1.0e-11$, which is depicted in Table C.1. Therefore, we can choose any small value of the disturbance level $\gamma > 0.1532178$. It should be noted that for any value of $\gamma < \gamma_{\min}$, the solution of GARE (C.7) does not exist and the H_∞-control cannot be designed.

Bibliography

[Abouheaf et al. (2013)] M. Abouheaf, F. Lewis, S. Haesaert, R. Babuska and K. Vamvoudakis, "Multi-agent discrete-time graphical games: interactive Nash equilibrium and value iteration solution", *American Control Conference (ACC)*, Washington, DC, USA, pp. 4189–4195, June, 2013.

[Afanasiev et al. (2013)] V.N. Afanasi'ev, V.B. Kolmanovskii, V.R. Nosov, Mathematical Theory of Control Systems Design, Springer Science & Business Media, vol. 341, 2013.

[Ahmed and Mukaidani (2016)] M. Ahmed and H. Mukaidani, "H_∞-constrained incentive Stackelberg game for discrete-time systems with multiple non-cooperative followers", *6th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys)*, Tokyo, Japan, September, 2016; *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 262–267, 2016.

[Ahmed et al. (2017a)] M. Ahmed, H. Mukaidani and T. Shima, "H_∞-constrained incentive Stackelberg games for discrete-time stochastic systems with multiple followers", *IET Control Theory & Applications*, vol. 11, no. 15, pp. 2475–2485, 2017.

[Ahmed et al. (2017b)] M. Ahmed, H. Mukaidani and T. Shima, "Infinite-horizon multi-leader-follower incentive stackelberg games for linear stochastic systems with H_∞ constraint", *56th Annual Conference of the Society of Instrument and Control Engineers (SICE)*, Kanazawa, Japan, pp. 1202–1207, September, 2017.

[Ahmed et al. (2018)] M. Ahmed, H. Mukaidani and T. Shima, "H_∞ constraint Pareto optimal strategy for stochastic LPV systems", *International Game Theory Review*, vol. 20, no. 2, pp. 1750031-1–1750031-20, 2018.

[Anderson and Moore (1989)] B.D.O. Anderson and J.B. Moore, "Optimal Control-Linear Quadratic Methods", Prentice-Hall, New York, 1989.

[Apkarian et al. (1995)] P. Apkarian, P. Gahinet, G. Becker, "Self-scheduling H_∞ control of linear parameter-varying systems: a design example", *Automatica*, vol. 31, no. 9, pp. 1251–1261, 1995.

[Arapostathis et al. (2010)] A. Arapostathis and V.S. Borkar, "Uniform properties of controlled diffusions and applications to optimal control", *SIAM Journal on Control and Optimization*, vol. 48, no. 7, pp. 4181–4223, 2010.

[Arrow et al. (1953)] K.J. Arrow, E.W. Barankin, D. Blackwell, H.W. Kuhn, A.W. Tucker, "Admissible Points of Convex Sets", Contributions to the Theory of Games, Princeton University Press, USA, 1953.

[Athans (1971)] M. Athans, "The role and use of the stochastic linear-quadratic-Gaussian problem in control system design", *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 529–552, 1971.

[Başar and Bernhard (2008)] T. Başar, P. Bernhard, "$H\infty$ Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach", Springer Science & Business Media, Birkhäuser, Boston, USA, 2008.

[Basar and Olsder (1980)] T. Başar and G. J. Olsder, "Team-optimal closed-loop Stackelberg strategies in hierarchical control problems", *Automatica*, vol. 16, no. 4, pp. 409–414, 1980.

[Başar and Olsder (1999)] T. Başar and G.J. Olsder, "Dynamic Noncooperative Game Theory", SIAM Series in Classics in Applied Mathematics, Philadelphia, Pennsylvania, USA, 1999.

[Basar and Selbuz (1979)] T. Basar and H. Selbuz, "Closed-loop Stackelberg strategies with applications in the optimal control of multi-level systems", *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 166–179, 1979.

[Boyd et al. (1994)] S. Boyd, L.E. Ghaoui, E. Feron and V. Balakrishnan, "Linear Matrix Inequalities in System and Control Theory", SIAM Studies in Applied Mathematics, vol. 15, Philadelphia, Pennsylvania, USA, 1994.

[Chen and Zhang (2004)] B.S. Chen, W. Zhang, "Stochastic H_2/H_∞ control with state-dependent noise", *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 45–57, 2004.

[Chen and Cruz (1972)] C.I. Chen and J.B. Cruz, "Stackelberg solution for two-person games with biased information patterns", *IEEE Transactions on Automatic Control*, vol. 17, no. 6, pp. 791–798, 1972.

[Dragan et al. (2006)] V. Dragan, T. Morozan and A.M. Stoica, "Mathematical Methods in Robust Control of Linear Stochastic Systems", Springer, New York, USA, 2006.

[Engwerda (2005)] J.C. Engwerda, "LQ Dynamic Optimization and Differential Games", John Wiley & Sons, West Sussex, England, 2005.

[Gao and Shi (2013)] Z. Gao and X. Shi, "Observer-based controller design for stochastic descriptor systems with Brownian motions", *Automatica*, vol. 49, no. 7, pp. 2229–2235, 2013.

[Gardner and Cruz (1978)] B.F. Gardner and J.B. Cruz, "Feedback Stackelberg strategy for M-level hierarchical games", *IEEE Transactions on Automatic Control*, vol. 23, no. 3, pp. 489–491, 1978.

[Ghaoui and Rami (1996)] L.E. Ghaoui and M.A. Rami, "Robust state-feedback stabilization of jump linear systems via LMIs", *International Journal of Robust and Nonlinear Control*, vol. 6, no. 9-10, pp. 1015–1022, 1996.

[Higham (2001)] D.J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations", *SIAM review*, vol. 43, no. 3, pp. 525–546, 2001.

[Henderson and Searle (1981)] H.V. Henderson and S.R. Searle, "The vec-permutation matrix, the vec operator and Kronecker products: A review", *Linear and multilinear algebra*, vol. 9, no. 4, pp. 271–288, 1981.

[Hinrichsen and Pritchard (1998)] D. Hinrichsen and A.J. Pritchard, "Stochastic H_∞", *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1504–1538, 1998.

[Ho et al. (1982)] Y.C. Ho, P.B. Luh and G.J. Olsder, "A control-theoretic view on incentives", *Automatica*, vol. 18, no. 2, pp. 167–179, 1982.

[Huang et al. (2008)] Y. Huang, W. Zhang, G. Feng, "Infinite horizon H_2/H_∞ control for stochastic systems with Markovian jumps", *Automatica*, vol. 44, no. 3, pp. 857–863, 2008.

[Isaacs (1999)] R. Isaacs, "Differential Games", Dover, New York, USA, 1999.

[Ishida and Shimemura (1983)] T. Ishida and E. Shimemura, "Three-level incentive strategies in differential games", *International Journal of Control*, vol. 38, no. 6, pp. 1135–1148, 1983.

[Kalman (1960)] R.E. Kalman, "Contributions to the theory of optimal control", *Bol. Soc. Mat. Mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[Ku and Wu (2015)] C.C. Ku, C.I. Wu, "Gain-scheduled H_∞ control for linear parameter varying stochastic systems", *Journal of Dynamic Systems Measurement and Control*, vol. 137, no. 11, pp. 111012-1–12, 2015.

[Lewis (1986)] F.L. Lewis, "Optimal Control", John Wiley & Sons, New York, USA, 1986.

[Lewis et al. (2013)] F.L. Lewis, H. Zhang, K.H-. Movric and A. Das, "Cooperative Control of Multi-agent Systems: Optimal and Adaptive Design Approaches", Springer Science & Business Media, London, UK, 2013.

[Li et al. (2002)] M. Li, J.B. Cruz, M.A. Simaan, "An Approach to Discrete-time Incentive Feedback Stackelberg Games", *IEEE Transaction Systems, Man, and Cybernetics*, vol. 32, no. 4, pp. 10–24, 2002.

[Liu et al. (2007)] L. Liu, X. Wei and X. Liu, "LPV Control for the Air Path System of Diesel Engines", *IEEE International Conference on Control and Automation*, Guangzhou, China, May 30–June 1, pp. 873–878, 2007.

[Luo et al. (2016)] S. Luo, Y. Sun, Y. Ji and D. Zhao, "Stackelberg game based incentive mechanisms for multiple collaborative tasks in mobile crowdsourcing", *Mobile Networks and Applications*, vol. 21, no. 3, pp. 506–522, 2016.

[Mahmoud (2002)] M.S. Mahmoud, "Discrete-time systems with linear parameter-varying: stability and H_∞-filtering", *Journal of mathematical analysis and applications*, vol. 269, no. 1, pp. 369–381, 2002.

[Mao (1997)] X. Mao, Stochastic Differential Equations and Applications, Horwood, Chichester, 1997.

[Medanic (1978)] J.V. Medanic, "Closed-loop Stackelberg Strategies in Linear-Quadratic Problems", *IEEE Transactions on Automatic Control*, vol. 23, no. 4, pp. 632–637, 1978.

[Mizukami and Wu (1987)] K, Mizukami, H. Wu, "Two-level incentive Stackelberg strategies in LQ differential games with Two noncooperative leaders and one follower", *Transactions of the Society Instrument and Control Engineers*, vol. 23, no. 6, 1987.

[Mizukami and Wu (1988)] K, Mizukami, H. Wu, "Incentive Stackelberg strategies in linear quadratic differential games with two noncooperative followers", *System Modelling and Optimization*, Lecture Notes in Control and Information Sciences, Springer, Berlin, Heidelberg, vol. 113, pp. 436–445, 1988.

[Moheimani (1996)] S.O.R. Moheimani and I.R. Petersen, "Optimal guaranteed cost control of uncertain systems via static and dynamic output feedback", *Automatica*, vol. 32, no. 4, pp. 575–579, 1996.

[Mukaidani (2009)] H. Mukaidani, "Robust guaranteed cost control for uncertain stochastic systems with multiple decision makers", *Automatica*, vol. 45, no. 7, pp. 1758–1764, 2009.

[Mukaidani et al. (2011)] H. Mukaidani, H. Xu and V. Dragan, "Multi-objective decision-making problems for discrete-time stochastic systems with state- and disturbance-dependent noise", *50th IEEE Conference on Decision and Control (CDC) and European Control Conference (ECC)*, Orlando, Florida, USA, pp. 6388–6393, December, 2011.

[Mukaidani (2014)] H. Mukaidani, "Stackelberg strategy for discrete-time stochastic system and its application to H_2/H_∞ control", *American Control Conference (ACC)*, Portland, Oregon, USA, pp. 4488–4493, June, 2014.

[Mukaidani and Xu (2015a)] H. Mukaidani and H. Xu, "Stackelberg strategies for stochastic systems with multiple followers", *Automatica*, vol. 53, no. (March 2015), pp. 53–59, 2015.

[Mukaidani et al. (2015b)] H. Mukaidani, M. Ahmed and H. Xu, "Finite horizon H_∞ control for stochastic systems with multiple decision makers", *54th IEEE Conference on Decision and Control (CDC)*, Osaka, Japan, pp. 513–518, December, 2015.

[Mukaidani (2016)] H. Mukaidani, "Infinite-horizon team-optimal incentive Stackelberg games for linear stochastic systems", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 99, no. 9, pp. 1721–1725, 2016.

[Mukaidani (2017a)] H. Mukaidani, Gain-scheduled H_∞ constraint Pareo optimal strategy for stochastic LPV systems with multiple decision makers, *American Control Conference (ACC)*, Seattle, WA, USA, pp. 1097–1102, May, 2017.

[Mukaidani et al. (2017b)] H. Mukaidani, H. Xu, T. Shima and V. Dragan, “A stochastic multiple-leader-follower incentive Stackelberg strategy for Markov jump linear systems”, *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 250–255, 2017.

[Mukaidani et al. (2017c)] H. Mukaidani, M. Ahmed, T. Shima and H. Xu, “H_∞ constraint incentive Stackelberg games for discrete-time stochastic systems”, *American Control Conference (ACC)*, Seattle, WA, USA, pp. 5257–5262, May, 2017.

[Mukaidani et al. (2017d)] H. Mukaidani, T. Shima, M. Unno, H. Xu and V. Dragan, “Team-optimal incentive Stackelberg strategies for Markov jump linear stochastic systems with H_∞ constraint”, *20th IFAC World Congress*, Toulouse, France, July, 2017; *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 3780–3785, 2017.

[Mukaidani and Xu (2018)] H. Mukaidani and H. Xu, “Incentive Stackelberg games for stochastic linear systems with H_∞ constraint”, *IEEE Transactions on Cybernetics*, 2018. (to appear)

[Mukaidani et al. (2018)] H. Mukaidani, H. Xu and V. Dragan, “Static output-feedback incentive Stackelberg game for discrete-time Markov jump linear stochastic systems with external disturbance”, *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 701–706, 2018.

[Osborne (2004)] M.J. Osborne, “An Introduction to Game Theory”, vol. 3, no., 3, Oxford University Press, UK, 2004.

[Ortega 1990] J.M. Ortega, Numerical Analysis, A Second Course. Philadelphia: SIAM, New York, USA, 1990.

[Pan and Başar (1993)] Z. Pan and T. Başar, “H_∞-optimal control for singularly perturbed systems. Part I: Perfect state measurements”, *Automatica*, vol. 29, no. 2, pp. 401–423, 1993.

[Pan and Yong (1991)] L. Pan and J. Yong, “A differential game with multi- level of hierarchy”, *Journal of Mathematical Analysis and Applications*, vol. 161, no. 2, pp. 522–544, 1991.

[Peng (1990)] S. Peng, “A general stochastic maximum principle for optimal control problems”, *SIAM Journal on Control and Optimization*, vol. 28, no. 4, pp. 966–979, 1990.

[Rami and Zhou (2000)] M.A. Rami, X.Y. Zhou, "Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls", *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1131–1143, 2000.

[Rami et al. (2002)] M.A. Rami, X. Chen and X.Y. Zhou, "Discrete-time indefinite LQ control with state and control dependent noises", *Journal of Global Optimization*, vol. 23, no. 3-4, pp. 245–265, 2002.

[Rotondo (2015)] D. Rotondo, V. Puig, F. Nejjari, "Linear quadratic control of LPV systems using static and shifting specifications", *European Control Conference (ECC)*, Linz, Austria, pp. 3085–3090, July, 2015.

[Saksena and Cruz (1985)] V.R. Saksena and J.B. Cruz, "Optimal and near-optimal incentive strategies in the hierarchical control of Markov chains", *Automatica*, vol. 21, no. 2, pp. 181–191, 1985.

[Salman and Cruz (1981)] M.A. Salman and J.B. Cruz, "An incentive model of duopoly with government coordination", *Automatica*, vol. 17, no. 6, pp. 821–829, 1981.

[Salman and Cruz (1983)] M.A. Salman and J.B. Cruz, "Team-optimal closed-loop Stackelberg strategies for systems with slow and fast modest", *International Journal of Control*, vol. 37, no. 6, pp. 1401–1416, 1983.

[Sato (2011)] M. Sato, "Gain-scheduled output-feedback controllers depending solely on scheduling parameters via parameter-dependent Lyapunov functions", *Automatica*, vol. 47, no. 12, pp. 2786–2790, 2011.

[Shamma (1988)] J. Shamma, "Analysis and design of gain scheduled control systems", PhD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, 1988.

[Shen (2004)] A. Bressan, W. Shen, "Semi-cooperative strategies for differential games", *International Journal of Game Theory*, vol. 32, no. 4, pp. 561–593, 2004.

[Simaan et al. (1973)] M. Simaan and J.B. Cruz, "Additional aspects of the Stackelberg strategy in nonzero-sum games", *Journal of Optimization Theory and Applications*, vol. 11, no. 6, pp. 613–620, 1973.

[Starr and Ho (1969)] A.W. Starr and Y.C. Ho, "Nonzero-sum differential games", *Journal of Optimization Theory and Applications*, vol. 3, no. 3, pp. 184–206, 1969.

[Tolwinski (1981)] B. Tolwinski, "Closed-loop Stackelberg solution to a multistage linear-quadratic game", *Journal of Optimization Theory and Applications*, vol. 34, no. 4, pp. 485–501, 1981.

[Wonham (1968)] W.M. Wonham, "On a matrix Riccati equation of stochastic control", *SIAM Journal on Control*, vol. 6, no. 4, pp. 681–697, 1968.

[Yamamoto 1986] T. Yamamoto, "A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions", *Numerische Mathematik*, vol. 49, no. 2-3, pp. 203–220, 1986.

[Zheng and Basar (1982)] Y.P. Zheng and T. Basar, "Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach", *International Journal of Control*, vol. 35, no. 6, pp. 997–1011, 1982.

[Zhang and Chen (2004)] W. Zhang and B.S. Chen, "On stabilizability and exact observability of stochastic systems with their applications", *Automatica*, vol. 40, no. 1, pp. 87–94, 2004.

[Zhang et al. (2006)] W. Zhang, H. Zhang and B.S. Chen, "Stochastic H_2/H_∞ control with (x, u, v)-dependent noise: finite horizon case", *Automatica*, vol. 42, no. 11, pp. 1891–1898, 2006.

[Zhang et al. (2007)] W. Zhang, Y. Huang and H. Zhang, "Stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise" *Automatica*, vol. 43, no. 3, pp. 513–521, 2007.

[Zhang et al. (2008)] W. Zhang, Y. Huang and L. Xie, "Infinite horizon stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise", *Automatica*, vol. 44, no. 9, pp. 2306–2316, 2008.

[Zeng et al. (2017)] T. Zeng, D. Upadhyay and G. Zhu, "Linear quadratic air-path control for diesel engines with regenerative and assisted turbocharger", *56th Annual Conference on Decision and Control (CDC)*, Melbourne, Australia, pp. 232–237, December, 2017.

[Zheng et al. (1984)] Y.P. Zheng, T. Basar and J.B. Cruz, "Stackelberg strategies and incentives in multiperson deterministic decision problems", *IEEE Transactions Systems, Man, and Cybernetics*, vol. 14, no. 1, pp. 10–24, 1984.