NASH STRATEGIES FOR DYNAMIC NONCOOPERATIVE LINEAR QUADRATIC SEQUENTIAL GAMES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Dan Shen, B.S., M.S.

* * * * *

The Ohio State University

2006

Dissertation Committee:

Jose B. Cruz, Jr., Adviser

Andrea Serrani

Philip Schniter

Approved by

Adviser

Electrical and Computer
Engineering Graduate Program

© Copyright by

Dan Shen

2006

ABSTRACT

Sequential games are those in which players apply their strategies following a certain predefined order. The sequential game is a natural framework to address some real problems, such as the “Action-Reaction-Counteraction” paradigm used in military intelligence and advertising campaigns strategies of several competing firms in economics. This dissertation focuses on stability, adaptivity and disturbance attenuation analysis of state feedback Nash strategies for a class of noncooperative dynamic Linear Quadratic Sequential Games (LQSGs).

For LQSGs with finite planning horizons, state feedback Nash strategies are provided and their existence and uniqueness within the class of state feedback strategies are proved. When the planning horizon approaches infinity, it is proved that the feedback systems with the state feedback Nash strategies are uniformly asymptotically stable, given that the associated coupled discrete-time algebraic Riccati equations (DAREs) have a positive definite solution. Moreover, it is also proved that at least one positive definite solution to the coupled DAREs of a scalar LQSG exists.

When the parameters of the objective functions are not shared among the decision makers, an on-line adaptive scheme is provided for each player to estimate the actual control gain used by the other player. The convergence to Nash strategies is proved with the condition that the associated coupled DAREs have a unique positive definite solution. The requirement of Persistency of Excitation (PE) is satisfied by

two methods: reference-signal tracking and system uncertainties represented by small white noises.

The disturbance attenuation analysis of state feedback Nash strategies is based on a H_∞ -optimal control problem, which is converted into a zero-sum game. In this approach, the uncertainties are controlled by a virtual player which maximizes the cost functions that are being minimized by the real decision makers. We first derive, for finite-horizon LQSGs, state feedback Nash strategies with the optimal attenuation level $\sqrt{\hat{\lambda}} > 0$. We extend the approach to infinite-horizon LQSGs. A lower bound $\hat{\lambda}_\infty$ is provided in the sense that for every $\lambda < \hat{\lambda}_\infty$ the values of the LQSGs are unbounded. We also prove that the feedback system with state feedback Nash strategies is Bounded Input Bounded Output (BIBO) stable with respect to the disturbances.

To
my parents
Hansheng Shen
and
Jianqiu Xu
for letting me pursue my dream
for so long
so far away from home
&
To my wife
Yi Lu
and my daughter
Michelle Shen
for
giving me new dreams to pursue

ACKNOWLEDGMENTS

I would like to express my appreciation to my adviser, Professor Jose B. Cruz, Jr., for the source of knowledge, guidance, and support that he has been to me during my graduate study at the Ohio State University. His enlightening advice and encouragement helped me achieve my goals. I could not have reached this far without his help.

I also want to thank Professor Vadim Utkin, Professor Kevin M. Passino, Professor Philip Schniter, and Professor Andrea Serrani for serving in my Ph.D candidacy committee and/or dissertation committee, Dr. Genshe Chen for his help with my research and studies, and my teammates: Ming Li, Huihui Jiang, Mo Wei, Xiaohuan Tan, Dongxu Li, Xu Wang and Ziyi Sun at the Ohio State University for their cooperation. My thanks also go to all the students from control group for their friendship during my stay here. I wish them great success in their studies and future careers.

Finally, I would like to thank my wife, Yi Lu, for her dedications to our family and her patience in waiting for me to finish school and to get a real job. I also thank my 4-year-old daughter, Michelle Shen, for all the joy and happiness she brought to my life.

VITA

April 22, 1974 Born - Suzhou, China

July 1998 B.S. in Automation
Tsinghua University
Beijing, China

August 1998 - August 2000 System Engineer
Softbrain Software (Beijing) Co. LTD
Beijing, China

September 2000 - December 2003 Graduate Research Associate
Dept. of Elec. & Computer Eng.
The Ohio State University
Columbus, OH 43210

March 2003 M.S. in Electrical Engineering
The Ohio State University
Columbus, OH 43210

January 2004 - December 2004 DAGSI Fellowship Award
Dept. of Elec. & Computer Eng.
The Ohio State University
Columbus, OH 43210

January 2005 - August 2005 Graduate Research Associate
Dept. of Elec. & Computer Eng.
The Ohio State University
Columbus, OH 43210

September 2005 - March 2006Intern
Intelligent Automation, Inc.
Rockville, MD 20855

April 2006 - present Graduate Research Associate
Dept. of Elec. & Computer Eng.
The Ohio State University
Columbus, OH

July 2006 - present Intern
Intelligent Automation, Inc.
Rockville, MD 20855

PUBLICATIONS

Research Publications

D. Shen and J. B. Cruz, Jr., “Design for Linear Quadratic Games with Adaptive State Feedback Nash Strategies,” Submitted to *IEEE Transactions of Automatic Control*, 2006.

D. Shen and J. B. Cruz, Jr., “Nash Strategies for Dynamic Noncooperative Linear Quadratic Sequential Games,” Submitted to *IEEE Transactions of Automatic Control*, 2006.

D. Shen and J. B. Cruz, Jr., “An Improved Game Theory Based Approach to One Type of H-infinity Optimal Control Problems,” in *Proceedings of the 2006 American Control Conference*, Minneapolis, Minnesota, June 14-16, 2006.

D. Shen, G. Chen, J. B. Cruz, Jr., C. Kwan, et al., “Game Theoretic Approach to Threat Intent Prediction,” in *Proceedings of CCRTS 2006: the Command and Control Research and Technology Symposium*, San Diego, July, 2006.

G. Chen, **D. Shen**, C. Kwan, J. B. Cruz, Jr., M. Kruger, “Game Theoretic Approach to Threat Prediction and Situation Awareness,” in *Proceedings of the 9th International Conference on Information Fusion*, Florence, Italy, July, 2006.

D. Shen and J. B. Cruz, Jr., “Encoding Strategy for Maximum Noise Tolerance Bidirectional Associative Memory,” in *IEEE Transactions of Neural Networks*, Vol. 16, No. 2, 2005.

D. Shen, J. B. Cruz, Jr., G. Chen, and C. Kwan, “A Game Theoretic Approach to Mission Planning For Multiple Aerial Platforms,” in *Proceedings of AIAA infotech@aerospace*, Arlington, VA., Sept. 26-29, 2005.

G. Chen, **D. Shen**, J. B. Cruz, Jr., C. Kwan, “A Novel Cooperative Path Planning for Multiple Aerospace Vehicles,” in *Proceedings of AIAA infotech@aerospace*, Arlington, VA., Sept. 26-29, 2005.

G. Chen, C. Kwan, **D. Shen**, J. B. Cruz, Jr., and A. Vannevel, “Adaptive Cooperative Path Planning for Multiple Platforms,” in *Proceedings of IEEE Robio 2005*, June 29-July 3, Hong Kong and Macau, 2005.

J. B. Cruz, Jr., G. Chen, D. Garagic, X. Tan, D. Li, **D. Shen**, M. Wei, X. Wang, “Team Dynamics and Tactics for Mission Planning,” in *Proceedings of IEEE Conference on Decision and Control*, Dec., 2003.

FIELDS OF STUDY

Major Field: Electrical Engineering

TABLE OF CONTENTS

	Page
Abstract	ii
Dedication	iv
Acknowledgments	v
Vita	vi
List of Figures	xii
Chapters:	
1. Introduction	1
1.1 Literature Review	1
1.1.1 Literature in Economics	2
1.1.2 Literature in Control	5
1.2 Problem Formulation and Motivation	9
1.3 Organization of the Dissertation	11
2. Existence and Stability Analysis of State Feedback Nash Strategies for LQSGs	13
2.1 Dynamic Noncooperative Linear Quadratic Sequential Games with Finite Planning Horizons	14
2.1.1 State Feedback Nash Strategies for Two-player LQSGs	14
2.1.2 Open-loop Nash Strategies for Two-player LQSGs	19
2.1.3 Nash Strategies for M -player LQSGs	24
2.2 Noncooperative Infinite-Horizon Linear Quadratic Sequential Games: Sequentially Switched Systems	26

2.2.1	State Feedback Nash Strategies	26
2.2.2	Stability Analysis of State Feedback Nash Strategies	28
2.3	Analysis of the Coupled Riccati Equations in the Scalar LQSG Cases	30
2.4	Computer Simulations	34
2.4.1	2-order Systems	34
2.4.2	A Scalar System	35
2.5	Chapter Summary	46
3.	Adaptation Design for Linear Quadratic Games with State Feedback Nash Strategies	47
3.1	Introduction	47
3.2	One-side Adaptation Scheme	49
3.2.1	Adaptation Design	50
3.2.2	Parameter Convergence and Persistency of Excitation	53
3.2.3	Extension to LQSGs	56
3.2.4	Simulations and Discussions	60
3.3	Two-side Adaptation Scheme	66
3.3.1	Adaptation Design	66
3.3.2	Parameter Convergence and Persistency of Excitation	69
3.3.3	Extension to LQSGs	78
3.3.4	Simulations and Discussions	79
3.4	Chapter Summary	88
4.	Disturbance Attenuation Analysis of State Feedback Nash Strategy for LQSG	90
4.1	Introduction	90
4.2	H^∞ -optimal Control and Related Zero-sum Games	92
4.3	Disturbance Attenuation Analysis of LQSGs with State Feedback Nash Strategies	93
4.3.1	Zero-sum Games for Disturbance Attenuation Analysis of LQSGs	93
4.3.2	State Feedback Nash Strategies for Finite-Horizon LQSGs with Soft-Constrained Cost Functions	94
4.3.3	Disturbance Attenuation Analysis for Finite-Horizon LQSGs with State Feedback Nash Strategies	102
4.3.4	Disturbance Attenuation Analysis of State Feedback Nash Strategies for Infinite-Horizon LQSGs	105
4.4	Simulations	113
4.5	Chapter Summary	115

5.	Conclusions	129
5.1	Contributions	129
5.1.1	Existence and Stability Analysis	129
5.1.2	Adaptation Design and Convergence Analysis	130
5.1.3	Disturbance Attenuation Analysis	132
5.2	Future Directions	133
	Bibliography	135

LIST OF FIGURES

Figure	Page
2.1 Asymptotically stable feedback system with Nash Strategy	36
2.2 K_k^1 obtained from Definition 1, K_e^1 and K_o^1 exist	37
2.3 K_k^2 obtained from Definition 1, K_e^2 and K_o^2 exist	38
2.4 Unstable feedback system with Nash Strategy	39
2.5 K_k^1 , obtained from Definition 1, diverges	40
2.6 K_k^2 , obtained from Definition 1, diverges	41
2.7 Function f_e and f_o	42
2.8 At least one solution to scalar DAREs exists.	43
2.9 K_k^1 and K_k^2 when $N = 11$	44
2.10 The feedback system is stable	45
3.1 Indirect adaptive control design for one-side adaptation scheme . . .	51
3.2 One-side adaptation for a simultaneous game: reference signal tracking	61
3.3 One-side adaptation for a simultaneous game with small system noise	62
3.4 One-side adaptation for a LQSG: reference signal tracking case	64
3.5 One-side adaptation for a LQSG with small system noise	65

3.6	Indirect adaptive control design for two-side adaptation scheme . . .	67
3.7	Reaction-curve-like relation between function f^1 and f^2	71
3.8	Three typical cases for a scalar system	74
3.9	Two-side adaptation for a simultaneous game: reference signal tracking	81
3.10	Non-increasing distances in the two-side adaptation for a simultaneous game: reference signal tracking	82
3.11	Two-side adaptation for a simultaneous game with small system noise	83
3.12	Two-side adaptation for a LQSG: reference signal tracking case	85
3.13	Non-increasing distances in the two-side adaptation for a LQSG: ref- erence signal tracking	86
3.14	Two-side adaptation for a LQSG with small system noise	87
4.1	Ξ_k , when $\hat{\lambda}_\infty = 6.2928$	116
4.2	K_k^1 , obtained from (4.12), converges	117
4.3	Eigenvalues of K_k^1 . K_o^1 and K_e^1 exist, and both are positive definite. .	118
4.4	K_k^2 , obtained from (4.13), converges	119
4.5	Eigenvalues of K_k^2 . K_o^2 and K_e^2 exist, and both are positive definite. .	120
4.6	Λ_k , obtained from (4.14), converges	121
4.7	Eigenvalues of Λ_k . Λ_o and Λ_e exist, and both are positive definite. . .	122
4.8	State feedback Nash strategies for real decision makers (u_k^1, u_k^2) and disturbance (w_k)	123
4.9	Asymptotically stable closed-loop system with state feedback Nash Strategies for real decision makers and disturbance.	124
4.10	State feedback Nash strategies for real decision makers (u_k^1, u_k^2)	125

4.11	The closed-loop system is BIBO stable	126
4.12	Cost-to-Go V_k^1 and V_k^2 when $w_k \sim \mathcal{N}(0, 0.1)$	127
4.13	$w_k \sim \mathcal{N}(0, 0.1)$, $(V_1^1) + (V_1^2) < V_1^w$, where $V_k^w = \hat{\lambda}_\infty \sum_{i=k}^{N-1} w_i^\top w_i$. . .	128

CHAPTER 1

INTRODUCTION

This dissertation studies stability, adaptivity, and disturbance attenuation analysis of state feedback Nash Strategies for a class of noncooperative dynamic Linear Quadratic Sequential Games (LQSGs). The background of the dissertation research is introduced by reviewing some of the most relevant areas to our research in this chapter. Afterwards, the problem formulation and motivation are summarized. Finally, the dissertation organization is described.

1.1 Literature Review

Game theory is a methodology to study the behavior of decision makers in situations, in which every individual decision maker is not in complete control of other decision units entering into the environment. The idea of a game can be tracked back to the Babylonian Talmud which is a compilation of ancient law and tradition set down during the first five centuries A.D.. However, it was not until 1944 that the mathematical theory of games was invented by John von Neumann and Oskar Morgenstern [1]. In that paper, some seminal work in areas of game theory such as the notion of a cooperative game, transferable utility (TU), coalitional form and von Neumann-Morgenstern stable sets were developed in terms of two-person zero sum

games. After that, tremendous efforts have been devoted by researchers from economics and control areas to various game strategies under this mathematical game framework.

1.1.1 Literature in Economics

Game theory is widely used in economics to model the markets and pricing strategies. As a main part of economic theory, game theory links other three main branches: decision theory, general equilibrium theory and mechanism design theory together. Game theory has been researched for a long time by economists and it is still of significant interest.

Between 1950 and 1953, John Nash [2, 3, 4, 5] made seminal contributions to both non-cooperative game theory and bargaining theory. In [2] and [3], Nash proved the existence of a strategic equilibrium for non-cooperative games - the Nash equilibrium - and proposed the “Nash Program”, in which reduction from cooperative games to non-cooperative form was suggested. In [4] and [5], Nash founded the axiomatic bargaining theory, proved the existence of the Nash bargaining solution, and provided the first execution of the “Nash Program”. Because of his enormous impact, Nash shared the Nobel prize in 1994 with John C. Harsanyi and Reinhard Selten.

In 1951, a learning concept, fictitious play, was first introduced by G. W. Brown [6]. Within the learning scheme, each player presumes that her opponents are playing stable (possibly mixed) strategies. Each player starts with some initial beliefs and chooses a best response to those beliefs as a strategy in this round. Then, after observing their opponents’ actions, the players update their beliefs according to some learning rule (e.g. Q-learning or Bayes’ rule). The process is then repeated. It is

known [7] that if it converges, then the point of convergence is a Nash equilibrium of the game.

In 1952-1953, the notion of the Core as a general solution concept was developed by L. S. Shapley and D. B. Gillies [8]. The core is the set of allocations that cannot be improved upon by any coalition. In [9], Shapley provided a set of axioms, a solution concept that associates with each coalitional game a unique out-come, which is now known as the Shapley Value. Shapley [10] also showed that for the strictly competitive case, with future payoff discounted at a fixed rate, such games are determined and that they have optimal strategies that depend only on the game being played, not on the history or even on the date, i.e., the strategies are stationary.

In 1953, H. W. Kuhn [11] proposed the extensive form of games. The extensive form (also called a game tree) is a graphical representation of a game. It allows modelers to specify the exact order in which players have to make their decisions and to formulate assumptions about the information possessed by the players in all stages of the game. The extensive form provides information about the players, payoffs, strategies, and the order of moves. The game tree consists of nodes, at which players can take actions, connected by edges, which represent the actions that may be taken at the previous node. Each terminal node is labeled with the payoffs earned by each player if the game ends at that node.

In 1966, infinitely repeated games with incomplete information were introduced by R. J. Aumann and M. Maschler [12]. 15 years later, R. J. Aumann [13] proposed the ideas of applying the notion of an automaton to describe a player in a repeated game and of studying the interactive behavior of bounded players, who draw their strategies only from (perhaps different) computationally bounded sets, via a game

with an appropriately restricted set of strategies. These ideas have given birth to a large and growing literature. Following R. J. Aumann, the theory of automata is now being used to formulate the idea of bounded rationality in repeated games. Two of the first authors to take this approach were A. Neyman [14] and A. Rubinstein [15].

In 1972, the concept of an Evolutionarily Stable Strategy (ESS) was introduced to evolutionary game theory by John M. Smith [16]. ESS is a general one, meaning, roughly, a strategy that, when used by an entire population, is resistant to invasion by new (mutant) strategies, i.e., is stable with respect to evolutionary changes. An Evolutionarily Stable Strategy is always a Nash Equilibrium. The major impetus for the use of the ESS concept can be found in [17] by John M. Smith and George R. Price. The ESS concept has since found increasing uses within the economics and biology literatures.

In 1974, R. J. Aumann proposed the concept of a correlated equilibrium in [18]. Unlike Nash equilibria, which are the concept of equilibria formulated in independent strategies, the correlated equilibria were developed from the correlated strategies in noncooperative games. Based on the signals generated by the correlated devices and announced to each of the decision makers, players choose their actions according to the received private signals.

The concept of subgame perfect equilibria was introduced by Selton [19] to refine the Nash equilibrium. A subgame perfect equilibrium is an equilibrium such that players' strategies constitute a Nash equilibrium in every subgame of the original game. Backward induction may be used to find the solutions. In 1982, Kreps and

Wilson [20] extended the idea of a subgame perfect equilibrium to subgames in extensive form that begin at imperfect information sets. This extended idea of equilibrium is called sequential equilibria.

A mixed strategy is used in game theory to describe a strategy comprising possible actions and an associated probability distribution which corresponds to how frequently each action is chosen. The first modern formulation of a mixed strategy was first set by E. Borel in his four notes on strategy games between 1921 and 1927. Then in [2], Nash proved that if each player in an N -person game has a finite number of pure strategies then there exists at least one Nash equilibrium, which can be either pure strategy Nash equilibrium or mixed strategy Nash equilibrium. In 1990, Crawford [21] discusses mixed strategy Nash equilibrium when the players preferences do not satisfy the assumptions necessary to be represented by expected utility functions. For discontinuous games, Reny [22] has recently proved the existence of mixed or pure strategy Nash equilibrium for a large class of games: those with finite player sets, compact and convex strategy sets, and a condition of better reply security holds.

In recent years, some improvements and refinements can be found in [23, 24, 25, 26, 27].

1.1.2 Literature in Control

In the optimal control framework, only one controller and some noises are included for consideration. For the more applicable case with more than one controller, game theory is a proper tool to analyze the behaviors of such systems. To emphasize the dynamic aspects of game theory, decades of efforts have been devoted to the task of adapting the economic concept to a system control environment.

Saddle points appear in many optimal control problems. But they can be viewed as solutions of minimax problems also, which arises in differential game problems where two or more players compete trying to maximize their objective functions and minimize the one of the others. It was proved by John von Neumann [28] that the existence of a minimax is guaranteed under very weak conditions. With the help of the Mountain Pass Lemma [29] (A mountain pass is indeed a beautiful example of saddle point provided by Nature), minimax control has been used systematically for solving nonlinear differential problems.

Differential Games were developed by Rufus P. Isaacs in the internal Rand Corporation research memoranda RM-1391 (30 November 1954), RM-1399 (30 November 1954), RM-1411 (21 December 1954) and RM-1486 (25 March 1955), which formed the basis of his famous book, *Differential Games (A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization)*[30], ten years later. Differential Games concern the balance of “optimal” strategies applied by two opposing decision makers, who have conflicting goals in the dynamical system they are both trying to control. Differential Game has its origins in pursuit-evasion games in a military context, but now it plays an important role in robust controller design.

In 1930, H. von Stackelberg [31] originally proposed the leader-follower solution concept, also known as the Stackelberg model, for static economic competition. In game theory terms, the players of this game are a leader and a follower and they compete on quantity. The leader moves first, choosing a quantity. The follower observes the leader’s choice and then picks a quantity. In Stackelberg game, the leader must know that the follower observes his action, and the follower must have no means of committing to a future non-Stackelberg follower action.

Then in the 1970s, Chen and Cruz [32], and Simaan and Cruz [33, 34], presented dynamic versions of the Stackelberg solution concept in a control context. In [32], Stackelberg Strategies were extended to the case of dynamic games with biased information pattern — one of the players knows only his own cost function but the other player knows both cost functions. In [33], the properties of the Stackelberg strategies in dynamic nonzero-sum two-player games were investigated to obtain necessary and sufficient conditions for existence and uniqueness of equilibrium strategies in linear-quadratic games, in the open-loop case. Additional aspects when the players use controls that are functions of state and time were discussed in [34]. For the discrete-time linear quadratic case, Li, Cruz and Simaan [35] suggested an incentive state feedback Stackelberg strategy. The authors obtained the team-optimal state feedback Stackelberg strategy (with no delays) of an important class of discrete-time two-person nonzero-sum dynamic games characterized by linear state dynamics and quadratic cost functionals.

T. Başar and H. Selbuz [36] use a one step delay information sharing to obtain closed-loop Stackelberg solutions. They proved that, under certain conditions, the closed-loop Stackelberg strategies for the leader are linear and of the one-step memory type, while those of the follower can be realized in linear feedback form.

The theoretical results of simultaneous dynamic games in the normal form by numerous authors were unified in [37] by T. Başar and G. J. Olsder. The existence, the uniqueness and some stability analysis of dynamic games (Nash, zero-sum, Stackelberg, and Pursuit-Evasion) were discussed with perfect information and incomplete information sets.

T. Başar and P. Bernhard [38] also extended the zero-sum game theory to the H_∞ -optimal control area. In this extension, the plant has two types of inputs (control signal and disturbance) and two types of outputs (regulated and measured). The controller is designed (based on the measured outputs) in such a way that the effect of the disturbance on the regulated output is minimized.

Reinforcement learning has been successfully used to find optimal control policies for a single agent operating in a stationary environment, specifically a Markov decision process. But learning in multi-agent systems is challenging (see [39]) due to the fact that from an agent’s perspective the world is not stationary — the behaviors of the other agents may change as they also learn to better perform their tasks. However, this non-stationary scenario can be viewed as a game problem in which there is an underlying assumption that the players have similar adaptation and learning abilities. Therefore, it seems promising to apply the relevant results from game theory to multi-agent reinforcement learning.

Q-learning [40], a recent form of reinforcement learning algorithm that does not need a model of its environment and can be used on-line, is widely and successfully used in Markov (Stochastic) Games [10], which is the extended version of the single agent Markov Decision Process (MDP). Markov games include multiple agents whose actions all impact the resulting rewards and next state. From the perspective of game theory, Markov games can be viewed also as an extension of matrix games. Such a model emphasizes the difficulty of finding optimal behavior in stochastic games, since optimal behaviors are dependent on the behaviors of the other agents, and vice versa.

The control concept of stability, usually used in the analysis of a closed-loop system associated with equilibrium points, can be applied to Nash equilibria also. A

Nash equilibrium for a mixed strategy game is stable if a small change (specifically, an infinitesimal change) in probabilities for one player leads to a situation where two conditions hold: 1) the player who did not change has no better strategy in the new circumstance; and 2) the player who did change is now playing with a strictly worse strategy. In this scenario, the player with the small change in his mixed-strategy will return immediately to the Nash equilibrium. Then the equilibrium is said to be stable. If condition one does not hold then the equilibrium is unstable. If only condition one holds then there are likely to be an infinite number of optimal strategies for the player who changed. It was proved by John Nash that the latter situation could not arise in a range of well-defined games.

In practical applications of Nash equilibria, the associated stability is crucial since the mixed-strategy of each player is not perfectly known, but has to be inferred from statistical distribution of his actions in the game. In this case unstable equilibria are very unlikely to arise in practice, since any minor change will lead to a change in strategy and then breakdown of the equilibrium.

1.2 Problem Formulation and Motivation

A sequential game is one in which players apply their strategies following a certain predefined order, and in which at least some players can observe the moves of other players who make decisions preceding them. The sequential game is a natural framework to address some real problems, such as the “Action-Reaction-Counteraction” [41, 42] paradigm used in military intelligence and advertising campaigns strategies of several competing firms in economics.

A sequential game was first introduced and researched by White in [43], but the results were based on an approximate ε -equilibria. Then in 1982, Kreps and Wilson [20] introduced a new criterion of “sequential equilibrium” for sequential games. Six years later, Breton, Alj and Haurie proposed a leader-follower one, “sequential Stackelberg Equilibria” in [44].

From the literature review, we find that sequential games are currently represented by game trees (the extensive form) and solved using the concept of rollback, or sub-game perfect equilibrium. The theoretical results for dynamic non-cooperative sequential game are very limited. It motivates us to do research on the theoretical analysis of the behavior of dynamic non-cooperative linear quadratic sequential games (LQSGs) in the normal form defined by

$$x_{k+1} = f_k(x_k, u_k^1, u_k^2, \dots, u_k^M; k) \quad (1.1)$$

$$y_k = g_k(I_k; k) \quad (1.2)$$

with the constraints (for $m \in [1, 2, \dots, M]$)

$$u_k^m = 0, \text{ if } m \neq (k \bmod M) \quad (1.3)$$

where, system states $x_k \in \mathcal{R}^n$, controllers $u_k^i \in \mathcal{R}^{r_i}$, measurements $y_k \in \mathcal{R}^r$, f_k and g_k are the linear system function and output function respectively, and I_k is the information set at time k , $1 \leq i \leq M$.

The cost function of the m^{th} ($m \in [1, 2, \dots, M]$) player is defined by

$$J^m = \sum_{k=0}^N h_k^m(y_k, u_k^1, u_k^2, \dots, u_k^M; k) \quad (1.4)$$

where h_k^m is the quadratic cost function for player m at time k , N is the time horizon.

For this kind of LQSGs, the existence of state feedback Nash strategies needs to be explored. When the time horizon approaches infinity, the stability of the closed-loop should be investigated. Since the parameters of the objective functions may not be shared among the decision makers, an on-line adaptive scheme should be designed for each player to estimate the actual controller used by other players. Convergence of the on-line adaptation need to be investigated. The requirement of Persistency of Excitation should be considered as well. In the real environment, the disturbance is always there, so we need to do some research on the disturbance attenuation analysis of our state feedback Nash strategies, too.

1.3 Organization of the Dissertation

The dissertation is organized as follows. In Chapter 2, we consider a class of noncooperative dynamic Linear Quadratic Sequential Games (LQSGs). For LQSGs with finite planning horizons, we provide state feedback Nash strategies, and their existence and uniqueness within the class of state feedback strategies are proved. When the planning horizon approaches infinity, we prove that the feedback systems with the state feedback Nash strategies are uniformly asymptotically stable, given that the associated coupled DAREs have a positive definite solution. Finally we show that at least one positive definite solution for the coupled Riccati equations of a scalar LQSG exists.

Chapter 3 focus on the adaptivity of infinite-horizon linear quadratic games (including simultaneous and sequential games) with state feedback Nash strategies. We first explore the one-side adaptation scheme, in which only one decision maker does adaptation to “guess” the strategies of the other player, while his opponent has access

to both players' cost functions. The requirement of Persistency of Excitation [45] is well considered via reference-signal tracking and small system noises. Secondly, we extend our approach to two-side adaptation, in which both players have no access to the parameters in other player's cost function. Convergence is guaranteed given that feedback Nash strategies uniquely exist.

In Chapter 4, we use a zero-sum game theoretical approach to address the disturbance attenuation analysis of our state feedback Nash strategies for LQSGs with uncertainties or disturbances. We first derive, for finite-horizon LQSGs, state feedback Nash strategies with the optimal attenuation level $\sqrt{\hat{\lambda}} > 0$. Then we extend the approach to infinite-horizon LQSGs. We prove that the feedback system is Bounded Input Bounded Output (BIBO) stable with respect to the disturbances.

Chapter 5 gives a summary of the major contributions of this dissertation and provides an outline of future research directions.

CHAPTER 2

EXISTENCE AND STABILITY ANALYSIS OF STATE FEEDBACK NASH STRATEGIES FOR LQSGS

Sequential games are those in which players make moves following a certain pre-defined order. A sequential game was first introduced and researched by White in [43], but the results were based on an approximate ε -equilibria. Then in 1982, Kreps and Wilson [20] introduced a new criterion of “sequential equilibrium” for sequential games. Six years later, Breton, Alj and Haurie proposed a leader-follower one, “sequential Stackelberg Equilibria” in [44]. Sequential games can be represented by game trees (the extensive form of games) and solved using the concept of rollback, or sub-game perfect equilibrium, which was proposed by Selten [19].

These current tools lack the sophistication to study games in which the players have relatively large, or even infinitely large, planning horizons. In this chapter, we develop a technique for sequential games in the context of linear-quadratic formulations with perfect information structure. Although linear-quadratic games have been studied extensively by Başar and Olsder in [37], the combination of dynamic programming and linear-quadratic sequential games is seldom touched.

The remainder of this chapter is organized as follows: In Section 2.1, we derive state feedback and open-loop Nash strategies for Linear Quadratic Sequential Games

(LQSGs) with finite planning horizons. We investigate the existence and uniqueness of these Nash strategies. Then, in Section 2.2, we prove that the infinite-horizon non-cooperative dynamic LQSG sequentially switched systems with our state feedback Nash Strategies are uniformly asymptotically stable with the assumption that the associated coupled DAREs have a positive definite solution. In Section 2.3, we prove the existence of solutions to the coupled DAREs for the scalar cases. Numerical systems are simulated in Section 2.4 to illustrate the asymptotical stability of the system with state feedback Nash strategies. Finally, some concluding remarks are provided in Section 2.5.

2.1 Dynamic Noncooperative Linear Quadratic Sequential Games with Finite Planning Horizons

In this section, we consider Nash strategies for noncooperative dynamic LQSGs. Dynamic programming and Hamiltonian are used to deduce Nash strategies for the 2-player finite games. We also extend our approach to the M -player cases.

2.1.1 State Feedback Nash Strategies for Two-player LQSGs

Consider the LQSG with 2 players of the form

$$x_{k+1} = A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2 \quad (2.1)$$

with the constraints

$$\begin{cases} u_k^1 = 0, & \text{when } k \text{ is even} \\ u_k^2 = 0, & \text{when } k \text{ is odd} \end{cases} \quad (2.2)$$

where, $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}^{r_i}$, $1 \leq i \leq 2$, $0 \leq k \leq N-1$.

The cost functions of the players are defined by

$$J^1 = x_N^\top Q_N^1 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + u_k^{2\top} R_k^{12} u_k^2) \quad (2.3)$$

$$J^2 = x_N^\top Q_N^2 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^2 x_k + u_k^{1\top} R_k^{21} u_k^1 + u_k^{2\top} R_k^{22} u_k^2) \quad (2.4)$$

where all matrices are symmetric with $Q_k^i \geq 0$ and $R_k^{ij} > 0$ for $i = 1, 2$, $j = 1, 2$, and $k = 0, 1, \dots, N$.

The Nash strategy was proposed originally by Nash in [3]. For the sequential game defined above, $\pi^* = \{(u_k^{1*}, u_{k+1}^{2*}) | k \text{ is odd, and } k \leq N-1\}$ is a Nash equilibrium (here assume the player 1 and player 2 will apply their controllers at odd-valued and even-valued time index, respectively), if

$$J^1(\pi^*) \leq J^1(\pi^1)$$

$$J^2(\pi^*) \leq J^2(\pi^2)$$

for every strategy $\pi^1 = \{(u_k^1, u_{k+1}^{2*}) | k \text{ is odd, } k \leq N-1\}$ and $\pi^2 = \{(u_k^{1*}, u_{k+1}^2) | k \text{ is odd, and } k \leq N-1\}$.

For simplicity of notations, we denote the following two conditions as

$$\text{Condition 1} \text{ --- } k \text{ is even} \quad := \quad \mathcal{K}_e$$

$$\text{Condition 2} \text{ --- } k \text{ is odd} \quad := \quad \mathcal{K}_o$$

Definition 1 *We define*

$$L_k^1 = \begin{cases} 0 & , \quad \mathcal{K}_e \\ -[R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1]^{-1} B_k^{1\top} K_{k+1}^1 A_k & , \quad \mathcal{K}_o \end{cases} \quad (2.5)$$

and

$$L_k^2 = \begin{cases} -[R_k^{22} + B_k^{2\top} K_{k+1}^2 B_k^2]^{-1} B_k^{2\top} K_{k+1}^2 A_k & , \quad \mathcal{K}_e \\ 0 & , \quad \mathcal{K}_o \end{cases} \quad (2.6)$$

where K_k^1 and K_k^2 are coupled in the recursive format of the related DAREs.

$$K_k^1 = \begin{cases} Q_k^1 + L_k^{2\top} R_k^{12} L_k^2 + (A_k + B_k^2 L_k^2)^\top K_{k+1}^1 (A_k + B_k^2 L_k^2), & \mathcal{K}_e \\ Q_k^1 + L_k^{1\top} R_k^{11} L_k^1 + (A_k + B_k^1 L_k^1)^\top K_{k+1}^1 (A_k + B_k^1 L_k^1), & \mathcal{K}_o \end{cases} \quad (2.7)$$

and

$$K_k^2 = \begin{cases} Q_k^2 + L_k^{2\top} R_k^{22} L_k^2 + (A_k + B_k^2 L_k^2)^\top K_{k+1}^2 (A_k + B_k^2 L_k^2), & \mathcal{K}_e \\ Q_k^2 + L_k^{1\top} R_k^{21} L_k^1 + (A_k + B_k^1 L_k^1)^\top K_{k+1}^2 (A_k + B_k^1 L_k^1), & \mathcal{K}_o \end{cases} \quad (2.8)$$

with boundary conditions $K_N^i = Q_N^i$.

Lemma 1 $K_k^1 \geq 0$ and $K_k^2 \geq 0$, for $\forall k \in [1, 2, \dots, N]$.

Proof: Since $Q_k^i \geq 0$ and $R_k^{ij} > 0$ for $i = 1, 2$, $j = 1, 2$, and $k = 0, 1, \dots, N$, by induction, we can prove Lemma 1. ■

Lemma 2 For $\forall k \in [0, 1, \dots, N-1]$, we have

$$(R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1) > 0$$

$$(R_k^{22} + B_k^{2\top} K_{k+1}^2 B_k^2) > 0$$

Proof: From Lemma 1, we can obtain

$$(B_k^{1\top} K_{k+1}^1 B_k^1) \geq 0$$

In addition, $R_k^{11} > 0$. Then $(R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1) > 0$. Similarly, we can obtain $(R_k^{22} + B_k^{2\top} K_{k+1}^2 B_k^2) > 0$. ■

Theorem 1 For the system defined in (2.1) - (2.4) with perfect information structure, the unique state feedback Nash strategies are given by

$$u_k^{1*} := \arg_{u_k^1} \min J^1 = \gamma_k^1(x_k) = L_k^1 x_k \quad (2.9)$$

$$u_k^{2*} := \arg_{u_k^2} \min J^2 = \gamma_k^2(x_k) = L_k^2 x_k \quad (2.10)$$

where L_k^1 and L_k^2 are defined, respectively, in (2.5) and (2.6).

Proof: We assume the optimal strategies γ_k^1 and γ_k^2 exist. Then we define

$$G_k^1 = x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + \gamma_k^{2\top} R_k^{12} \gamma_k^2 + V_{k+1}^1 \quad (2.11)$$

$$G_k^2 = x_k^\top Q_k^2 x_k + u_k^{2\top} R_k^{22} u_k^2 + \gamma_k^{1\top} R_k^{21} \gamma_k^1 + V_{k+1}^2 \quad (2.12)$$

$$V_k^1 = \min_{u_k^1} G_k^1 \quad (2.13)$$

$$V_k^2 = \min_{u_k^2} G_k^2 \quad (2.14)$$

and $V_N^i(x_N) = x_N^\top Q_N^i x_N$. For player 1, when $k = N - 1$,

$$G_{N-1}^1 = x_{N-1}^\top Q_{N-1}^1 x_{N-1} + u_{N-1}^{1\top} R_{N-1}^{11} u_{N-1}^1 + \gamma_{N-1}^{2\top} R_{N-1}^{12} \gamma_{N-1}^2 + x_N^\top Q_N^1 x_N$$

and $x_N = A_{N-1} x_{N-1} + B_{N-1}^1 u_{N-1}^1 + B_{N-1}^2 \gamma_{N-1}^2$. Then

$$\begin{aligned} G_{N-1}^1 &= u_{N-1}^{1\top} (R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1) u_{N-1}^1 \\ &+ 2u_{N-1}^{1\top} B_{N-1}^{1\top} Q_N^1 (A_{N-1} x_{N-1} + B_{N-1}^2 \gamma_{N-1}^2) + x_{N-1}^\top A_{N-1}^\top Q_N^1 A_{N-1} x_{N-1} \\ &+ 2x_{N-1}^\top A_{N-1}^\top Q_N^1 B_{N-1}^2 \gamma_{N-1}^2 + \gamma_{N-1}^{2\top} B_{N-1}^{2\top} Q_N^1 B_{N-1}^2 \gamma_{N-1}^2 \end{aligned} \quad (2.15)$$

The gradient of G_{N-1}^1 along u_{N-1}^1 is

$$\nabla_{u_{N-1}^1} G_{N-1}^1 = 2(R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1) + 2B_{N-1}^{1\top} Q_N^1 (A_{N-1} x_{N-1} + B_{N-1}^2 \gamma_{N-1}^2)$$

Since $R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1 > 0$, G_{N-1}^1 is strictly convex in u_{N-1}^1 . Set $\nabla_{u_{N-1}^1} G_{N-1}^1 = \vec{0}$, we obtain the unique state feedback Nash strategies for Player 1 at time $k = N - 1$

$$\gamma_{N-1}^1 = -[R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1]^{-1} B_{N-1}^{1\top} Q_N^1 (A_{N-1} x_{N-1} + B_{N-1}^2 \gamma_{N-1}^2) \quad (2.16)$$

Similarly, we get γ_{N-1}^2 as a function of γ_{N-1}^1

$$\gamma_{N-1}^2 = -[R_{N-1}^{22} + B_{N-1}^{2\top} Q_N^2 B_{N-1}^2]^{-1} B_{N-1}^{2\top} Q_N^2 (A_{N-1} x_{N-1} + B_{N-1}^1 \gamma_{N-1}^1) \quad (2.17)$$

Apply the constraints in (2.2), when N is odd, $\gamma_{N-1}^1 = 0$ and

$$\gamma_{N-1}^2 = -[R_{N-1}^{22} + B_{N-1}^{2\top} Q_N^2 B_{N-1}^2]^{-1} B_{N-1}^{2\top} Q_N^2 A_{N-1} x_{N-1}$$

when N is even, $\gamma_{N-1}^2 = 0$ and

$$\gamma_{N-1}^1 = -[R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1]^{-1} B_{N-1}^{1\top} Q_N^1 A_{N-1} x_{N-1}$$

So the conclusion in this theorem is correct for time step $N - 1$. In this proof, we will prove, as a byproduct, that

$$V_k^1 = x_k^\top K_k^1 x_k \quad (2.18)$$

$$V_k^2 = x_k^\top K_k^2 x_k \quad (2.19)$$

which are true for time step $N - 1$ by setting $u_{N-1}^i = L_{N-1}^i x_{N-1}$ and $V_N^i = x_N^\top K_N^i x_N$ in (2.11) and (2.12). Next we assume that equations (2.9) - (2.10) and (2.18) - (2.19) are true for time step $k + 1$. W.L.O.G, let k be odd-valued. Then $k + 1$ is even.

$$\begin{aligned} u_{k+1}^{1*} &= \gamma_{k+1}^1(x_{k+1}) = L_{k+1}^1 x_{k+1} \\ u_{k+1}^{2*} &= \gamma_{k+1}^2(x_{k+1}) = L_{k+1}^2 x_{k+1} \end{aligned}$$

where

$$\begin{aligned} L_{k+1}^1 &= 0 \\ L_{k+1}^2 &= -[R_{k+1}^{22} + B_{k+1}^{2\top} K_{k+2}^2 B_{k+1}^2]^{-1} B_{k+1}^{2\top} K_{k+2}^2 A_k \end{aligned}$$

then, since k is odd

$$u_k^2 = 0 \quad (2.20)$$

So for player 1,

$$G_k^1 = x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + x_{k+1}^\top K_{k+1}^1 x_{k+1} \quad (2.21)$$

$$V_k^1 = \min_{u_k^1} G_k^1 \quad (2.22)$$

and $x_{k+1} = A_k x_k + B_k^1 u_k^1$. Then

$$G_k^1 = u_k^{1\top} (R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1) u_k^1 + 2u_k^{1\top} B_k^{1\top} K_{k+1}^1 A_k x_k + x_k^\top A_k^\top K_{k+1}^1 A_k x_k$$

By Lemma 2, $R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1 > 0$, so G_k^1 is a strictly convex function of u_k^1 . Setting $\nabla_{u_k^1} G_k^1 = \vec{0}$, we obtain the unique state feedback Nash strategies for player 1 at time k

$$\gamma_k^1 = -[R_k^{11} + B_k^{1\top} K_{k+1}^1 B_k^1]^{-1} B_k^{1\top} K_{k+1}^1 A_k x_k = L_k^1 x_k$$

By setting $u_k^1 = L_k^1 x_k$ in (2.21), we obtain

$$V_k^1 = x_k^\top [Q_k^1 + L_k^{1\top} R_k^{11} L_k^1 + (A_k + B_k^1 L_k^1)^\top K_{k+1}^1 (A_k + B_k^1 L_k^1)] x_k$$

so, $V_k^1 = x_k^\top K_k^1 x_k$. Similarly, from (2.14) we get

$$V_k^2 = x_k^\top [Q_k^2 + L_k^{1\top} R_k^{21} L_k^1 + (A_k + B_k^1 L_k^1)^\top K_{k+1}^2 (A_k + B_k^1 L_k^1)] x_k = x_k^\top K_k^2 x_k$$

By induction, we complete our proof. ■

Remark 1 By Lemma 1 and Lemma 2, the control strategies given by (2.9)-(2.10) is a valid pair of Nash state feedback strategies for the Noncooperative Finite Sequential Game defined in (2.1)-(2.4). Uniqueness is guaranteed since for each time k , the G_k^i is strictly convex in u_k^i .

2.1.2 Open-loop Nash Strategies for Two-player LQSGs

In the open-loop information structure, no player has access to system states $\{x_k\}, k = 1, 2, \dots, N$. The open-loop Nash strategy may be a function of the initial state x_0 . In this section, we use a Hamiltonian based method (the technique will be used in Chapter 4) instead of the Isaacs' equation [46](Section 2.3), which was used in section 2.1.1, to derive open-loop Nash strategies for two-person LQSGs.

For the same system defined in (2.1)-(2.4), we define the new cost functions for each player,

$$Z^1 = \frac{1}{2} \sum_{k=0}^{N-1} (x_{k+1}^\top Q_{k+1}^1 x_{k+1} + u_k^{1\top} R_k^{11} u_k^1 + u_k^{2\top} R_k^{12} u_k^2) \quad (2.23)$$

$$Z^2 = \frac{1}{2} \sum_{k=0}^{N-1} (x_{k+1}^\top Q_{k+1}^2 x_{k+1} + u_k^{1\top} R_k^{21} u_k^1 + u_k^{2\top} R_k^{22} u_k^2) \quad (2.24)$$

Compare (2.3)-(2.4) with (2.23)-(2.24), we can find that

$$J^i = 2Z^i + x_0^\top Q_0^i x_0$$

So the optimal strategies for (2.23)-(2.24) is also optimal in the sense of (2.3)-(2.4), and *vice versa*.

The Hamiltonian of each player, under the new object functions (2.23)-(2.24), is

$$\begin{aligned} H_k^1 &= \frac{1}{2} (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2)^\top Q_{k+1}^1 (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) \\ &+ \frac{1}{2} u_k^{1\top} R_k^{11} u_k^1 + \frac{1}{2} u_k^{2\top} R_k^{12} u_k^2 + p_{k+1}^\top (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) \end{aligned} \quad (2.25)$$

$$\begin{aligned} H_k^2 &= \frac{1}{2} (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2)^\top Q_{k+1}^2 (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) \\ &+ \frac{1}{2} u_k^{1\top} R_k^{21} u_k^1 + \frac{1}{2} u_k^{2\top} R_k^{22} u_k^2 + p_{k+1}^{2\top} (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) \end{aligned} \quad (2.26)$$

Since $B_k^{i\top} Q_{k+1}^i B_k^i + R_k^{ii} > 0$, H_k^i is strictly convex in u_k^i . The minimization of H_k^i over u_k^i yields the unique relation

$$\frac{\partial H_k^i}{\partial u_k^i} = B_k^{i\top} [Q_{k+1}^i (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) + p_{k+1}^i] + R_k^{ii} u_k^i = \vec{0} \quad (2.27)$$

The costate p_k^i should satisfy the following equation

$$p_k^i = \frac{\partial H_k^i}{\partial x_k^i} = A_k^\top [Q_{k+1}^i (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) + p_{k+1}^i] \quad (2.28)$$

with boundary conditions $p_N^i = \vec{0}$.

Theorem 2 For a two-person LQSG defined in (2.1)-(2.2) with objective functions (2.23)-(2.24), let Λ_k and M_k^i ($i = 1, 2$) be matrices of appropriate dimensions, defined through

$$\Lambda_k = \begin{cases} I + B_k^2(R_k^{22})^{-1}B_k^{2\top}M_{k+1}^2 & , \quad \mathcal{K}_e \\ I + B_k^1(R_k^{11})^{-1}B_k^{1\top}M_{k+1}^1 & , \quad \mathcal{K}_o \end{cases} \quad (2.29)$$

$$M_k^i = \begin{cases} Q_k^i + A_k^\top M_{k+1}^i \Lambda_k^{-1} A_k & , \quad k = 0, 1, \dots, N-1 \\ Q_N^i & , \quad k = N \end{cases} \quad (2.30)$$

Then,

(i) the matrices Λ_k , recursively defined in (2.29), are invertible.

(ii) the unique open-loop Nash strategies are

$$\gamma_k^1 = u_k^{1*} = \begin{cases} 0 & , \quad \mathcal{K}_e \\ -(R_k^{11})^{-1}B_k^{1\top}M_{k+1}^1\Lambda_k^{-1}A_kx_k^* & , \quad \mathcal{K}_o \end{cases} \quad (2.31)$$

and

$$\gamma_k^2 = u_k^{2*} = \begin{cases} -(R_k^{22})^{-1}B_k^{2\top}M_{k+1}^2\Lambda_k^{-1}A_kx_k^* & , \quad \mathcal{K}_e \\ 0 & , \quad \mathcal{K}_o \end{cases} \quad (2.32)$$

where

$$x_{k+1}^* = A_kx_k^* + B_k^1\gamma_k^1 + B_k^2\gamma_k^2, \quad x_0^* = x_0 \quad (2.33)$$

Proof: When $k = N-1$, if $N-1$ is odd-valued, then $\gamma_{N-1}^2 = \vec{0}$ (from constraints (2.2)) and $\gamma_{N-1}^1 = -(R_{N-1}^{11})^{-1}B_{N-1}^{1\top}(Q_N^1x_N^* + p_N^1)$ (from (2.27)). Since $M_N^i = Q_N^i$ and $p_N^i = \vec{0}$, the optimal control at time $N-1$ for player 1 is

$$\gamma_{N-1}^1 = -(R_{N-1}^{11})^{-1}B_{N-1}^{1\top}M_N^1x_N^*$$

and if both sides of the above equation are multiplied by B_{N-1}^1 , we can obtain

$$x_N^* - A_{N-1}x_{N-1}^* = (I - \Lambda_{N-1})x_N^*$$

by using (2.29) and (2.33). Since the pair of strategies $(\gamma_{N-1}^1, \gamma_{N-1}^2)$ is the unique solution to (2.27) with constraints (2.2) at time $N-1$, there necessarily exist a unique relationship between x_{N-1}^* and x_N^* , implying that the matrix Λ_{N-1} should be invertible. So,

$$x_N^* = \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}^* \quad (2.34)$$

Then

$$\gamma_{N-1}^1 = -(R_{N-1}^{11})^{-1} B_{N-1}^{1\top} M_N^1 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}^* \quad (2.35)$$

Similarly, if $N-1$ is even, then $\gamma_{N-1}^1 = \vec{0}$ and

$$\gamma_{N-1}^2 = -(R_{N-1}^{22})^{-1} B_{N-1}^{2\top} M_N^2 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}^* \quad (2.36)$$

By (2.28) we can obtain

$$p_{N-1}^i = A_{N-1}^\top (Q_N^i x_N^* + p_N^i) = A_{N-1}^\top M_N^i x_N^* \quad (2.37)$$

To use the induction argument, we assume that (2.31)-(2.32),

$$x_{k+1}^* = \Lambda_k^{-1} A_k x_k^* \quad (2.38)$$

and

$$p_k^i = A_k^\top M_{k+1}^i x_{k+1}^* \quad (2.39)$$

hold for $k = l$. It is already verified for $l = N-1$. When $l-1$ is even, $\gamma_{l-1}^1 = \vec{0}$ and

$$\begin{aligned} \gamma_{l-1}^2 &= -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} [Q_l^2 x_l^* + p_l^2] \\ &= -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} [Q_l^2 x_l^* + A_l^\top M_{l+1}^2 x_{l+1}^*] \\ &= -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} [Q_l^2 x_l^* + A_l^\top M_{l+1}^2 \Lambda_l^{-1} A_l x_l^*] \\ &= -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} [Q_l^2 + A_l^\top M_{l+1}^2 \Lambda_l^{-1} A_l] x_l^* \\ &= -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} M_l^2 x_l^* \end{aligned}$$

If both sides of the above equation are multiplied by B_{l-1}^2 , we can obtain

$$x_l^* - A_{l-1}x_{l-1}^* = (I - \Lambda_{l-1})x_l^*$$

by using (2.29) and (2.33). Since the pair of strategies $(\gamma_{l-1}^1, \gamma_{l-1}^2)$ is the unique solution to (2.27) with constraints (2.2) at time $l-1$, there necessarily exist a unique relationship between x_{l-1}^* and x_l^* , implying that the matrix Λ_{l-1} should be invertible.

$$x_l^* = \Lambda_{l-1}^{-1} A_{l-1} x_{l-1}^* \quad (2.40)$$

Then

$$\gamma_{l-1}^2 = -(R_{l-1}^{22})^{-1} B_{l-1}^{2\top} M_l^2 \Lambda_{l-1}^{-1} A_{l-1} x_{l-1}^* \quad (2.41)$$

Similarly, if $l-1$ is even, then $\gamma_{l-1}^1 = \vec{0}$ and

$$\gamma_{l-1}^1 = -(R_{l-1}^{11})^{-1} B_{l-1}^{1\top} M_l^1 \Lambda_{l-1}^{-1} A_{l-1} x_{l-1}^* \quad (2.42)$$

By (2.28) we can obtain

$$p_{l-1}^i = A_{l-1}^\top (Q_l^i x_l^* + p_l^i) = A_{l-1}^\top M_l^i x_l^* \quad (2.43)$$

because $Q_l^i x_l^* + p_l^i = Q_l^i x_l^* + A_l^\top M_{l+1}^i x_{l+1}^* = Q_l^i x_l^* + A_l^\top M_{l+1}^i \Lambda_l^{-1} A_l x_l^* = M_l^i x_l^*$.

From (2.40) - (2.43), we prove that (2.31)-(2.32), (2.38), and (2.39) hold for $k = l-1$. By induction, we complete our proof. ■

Remark 2 *We have obtained the open-loop Nash strategies for LQSGs by using a Hamiltonian based method. The format of the controller design equations in (2.31)-(2.31) is totally different from the state feedback Nash strategies in (2.9)-(2.10). We can not obtain the open-loop Nash strategies by simply replacing x_k with x_k^* in (2.9)-(2.10). We notice that R_k^{12} and R_k^{21} do not appear in the open-loop Nash strategies.*

The reason is that in state feedback perfect information structure, we have to take into account the possible dependence of u_k^i on x_k . So, in the state feedback Nash strategies case, the costate dynamics (2.39) becomes

$$\begin{aligned} p_k^1 &= \frac{\partial H_k^1}{\partial x_k^1} = (A_k + B_k^2 L_k^2)^\top [Q_{k+1}^1 (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) + p_{k+1}^1] + L_k^{2\top} R_k^{12} u_k^2 \\ p_k^2 &= \frac{\partial H_k^2}{\partial x_k^2} = (A_k + B_k^1 L_k^1)^\top [Q_{k+1}^2 (A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2) + p_{k+1}^2] + L_k^{1\top} R_k^{21} u_k^1 \end{aligned}$$

where

$$\frac{\partial u_k^i}{\partial x_k^i} = L_k^{i\top}$$

Then, R_k^{21} and R_k^{12} enter the state feedback Nash Strategies.

2.1.3 Nash Strategies for M -player LQSGs

It is straightforward to extend the results to M -player cases. Consider the LQSG with M players of the form

$$x_{k+1} = A_k x_k + \sum_{m=1}^M B_k^m u_k^m \quad (2.44)$$

with the constraints (for $k \in [0, 1, \dots, N-1]$ and $m \in [1, 2, \dots, M]$)

$$u_k^m = 0, \text{ if } m \neq (k \bmod M) \quad (2.45)$$

where, $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}^{r_i}$, $1 \leq i \leq M$, $0 \leq k \leq N-1$. For simplicity of notations, we denote the following four conditions as

$$\text{Condition 3} \text{ --- } m \neq (k \bmod M) \quad := \quad \bar{\mathcal{M}}$$

$$\text{Condition 4} \text{ --- } m = (k \bmod M) \quad := \quad \mathcal{M}$$

$$\text{Condition 5} \text{ --- } k \neq N \quad := \quad \bar{\mathcal{N}}$$

$$\text{Condition 6} \text{ --- } k = N \quad := \quad \mathcal{N}$$

The cost function of the $m^{\text{th}}(m \in [1, 2, \dots, M])$ player is defined by

$$J^m = x_N^\top Q_N^m x_N + \sum_{i=0}^{N-1} \left(x_i^\top Q_i^m x_i + \sum_{p=1}^M u_i^{p\top} R_i^{mp} u_i^p \right) \quad (2.46)$$

where all matrices are symmetric with $Q_k^i \geq 0$ and $R_k^{ij} > 0$ for $i = 1, 2, \dots, M$, $j = 1, 2, \dots, M$, and $k = 0, 1, \dots, N$

By dynamic programming and induction, we obtain the state feedback Nash strategy for the $m^{\text{th}}(m \in [1, 2, \dots, M])$ player in the finite sequential game defined in (2.44)-(2.46). It is specified by

$$u_k^{m*} := \arg_{u_k^m} \min J^m = \gamma_k^m(x_k) = L_k^m x_k \quad (2.47)$$

$$L_k^m = \begin{cases} 0 & , \bar{\mathcal{M}} \\ -[\bar{R}^{mm}]^{-1} B_k^{m\top} K_{k+1}^m A_k & , \mathcal{M} \end{cases} \quad (2.48)$$

where $\bar{R}^{mm} = R_k^{mm} + B_k^{m\top} K_{k+1}^m B_k^m$.

$$K_k^m = \begin{cases} Q_k^m + L_k^{q\top} R_k^{mq} L_k^q + (\bar{A}_k^q)^\top K_{k+1}^m \bar{A}_k^q & , \bar{\mathcal{N}} \\ Q_N^m & , \mathcal{N} \end{cases} \quad (2.49)$$

where $q = (k \bmod M)$ and $\bar{A}_k^q = A_k + B_k^q L_k^q$.

Similarly, we have $K_k^m \geq 0$, for $\forall k \in [1, N]$ and $m \in [1, 2, \dots, M]$, and $(R_k^{mm} + B_k^{m\top} K_{k+1}^m B_k^m) > 0$, for $\forall k \in [0, N-1]$ and $m \in [1, 2, \dots, M]$.

We can also extend the two-person open-loop Nash strategies (2.31)-(2.32) to the M -player case. The open-loop Nash strategies for the system defined in (2.44)-(2.46) are

$$\gamma_k^m = u_k^{m*} = \begin{cases} 0 & , \bar{\mathcal{M}} \\ -(R_k^{mm})^{-1} B_k^{m\top} M_{k+1}^m \Lambda_k^{-1} A_k x_k^* & , \mathcal{M} \end{cases}$$

where

$$\Lambda_k = I + B_k^m (R_k^{mm})^{-1} B_k^{m\top} M_{k+1}^m, \quad m = (k \bmod M)$$

$$M_k^i = \begin{cases} Q_k^i + A_k^\top M_{k+1}^i \Lambda_k^{-1} A_k & , \quad k = 0, 1, \dots, N-1 \\ Q_N^i & , \quad k = N \end{cases}$$

and

$$x_{k+1}^* = A_k x_k^* + \sum_{i=1}^M B_k^i \gamma_k^i, \quad x_0^* = x_0$$

2.2 Noncooperative Infinite-Horizon Linear Quadratic Sequential Games: Sequentially Switched Systems

2.2.1 State Feedback Nash Strategies

For the LQSG defined in (2.44)-(2.46), when $N \rightarrow +\infty$, it becomes an infinite-horizon LQSG problem. If matrices $A_k, B_k^m, Q_k^m, R_k^{mp}$ are constant, the infinite-horizon sequential game is a sequentially switched system, which is a special case of the switched system in [47, 48]. Let us examine the 2-player cases defined as

$$x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2 \quad (2.50)$$

with the constraints

$$\begin{cases} u_k^1 = 0, & \text{when } k \text{ is even} \\ u_k^2 = 0, & \text{when } k \text{ is odd} \end{cases} \quad (2.51)$$

where, $x_k \in \mathcal{R}^n, u_k^i \in \mathcal{R}^{r_i}, 1 \leq i \leq 2, 0 \leq k < +\infty$.

The cost functions of the players are defined by

$$J^1 = \sum_{i=0}^{+\infty} (x_i^\top Q^1 x_i + u_i^{1\top} R^{11} u_i^1 + u_i^{2\top} R^{12} u_i^2) \quad (2.52)$$

and

$$J^2 = \sum_{i=0}^{+\infty} (x_i^\top Q^2 x_i + u_i^{1\top} R^{21} u_i^1 + u_i^{2\top} R^{22} u_i^2) \quad (2.53)$$

where all matrices are symmetric with $Q^i \geq 0$, (A, C^i) are observable (where $C^{i\top} C^i = Q^i$), and $R^{ij} > 0$ for $i = 1, 2$ and $j = 1, 2$.

Theorem 3 For K_k^i in Definition 1 with $A_k = A$, $B_k^i = B^i$, $Q_k^i = Q^i$ and $R_k^{ij} = R^{ij}$, if there exists K_e^i and K_o^i , and for any fixed even-valued k , $K_k^i \rightarrow K_e^i$ and for any fixed odd-valued k , $K_k^i \rightarrow K_o^i$ as $N \rightarrow \infty$, then the state feedback Nash strategies of the game defined in (2.50)-(2.53) are given by

$$u_k^{1*} := \arg_{u_k^1} \min J^1 = \gamma_k^1(x_k) = \bar{L}_k^1 x_k \quad (2.54)$$

$$u_k^{2*} := \arg_{u_k^2} \min J^2 = \gamma_k^2(x_k) = \bar{L}_k^2 x_k \quad (2.55)$$

where

$$\bar{L}_k^1 = \begin{cases} 0 & , \mathcal{K}_e \\ -[R^{11} + B^{1\top} K_e^1 B^1]^{-1} B^{1\top} K_e^1 A & , \mathcal{K}_o \end{cases} \quad (2.56)$$

$$\bar{L}_k^2 = \begin{cases} -[R^{22} + B^{2\top} K_o^2 B^2]^{-1} B^{2\top} K_o^2 A & , \mathcal{K}_e \\ 0 & , \mathcal{K}_o \end{cases} \quad (2.57)$$

Proof: Using the conclusion and the same proof method of Theorem 1, we obtain the state feedback Nash strategies by taking the limits of the Nash solutions for the finite planning horizon cases as $N \rightarrow \infty$. ■

Remark 3 From the conditions of Theorem 3, we obtain the coupled discrete-time Algebraic Riccati Equations (DAREs) associated with LQSGs

$$K_e^1 = Q^1 + L^{2\top} R^{12} L^2 + (A + B^2 L^2)^\top K_o^1 (A + B^2 L^2)$$

$$K_o^1 = Q^1 + L^{1\top} R^{11} L^1 + (A + B^1 L^1)^\top K_e^1 (A + B^1 L^1)$$

$$K_e^2 = Q^2 + L^{2\top} R^{22} L^2 + (A + B^2 L^2)^\top K_o^2 (A + B^2 L^2)$$

$$K_o^2 = Q^2 + L^{1\top} R^{21} L^1 + (A + B^1 L^1)^\top K_e^2 (A + B^1 L^1)$$

where

$$L^1 = -[R^{11} + B^{1\top} K_e^1 B^1]^{-1} B^{1\top} K_e^1 A$$

$$L^2 = -[R^{22} + B^{2\top} K_o^2 B^2]^{-1} B^{2\top} K_o^2 A$$

It is well known that there are no general existence conditions for the coupled DAREs associated with simultaneous games, see Başar and Olsder [37] (pp. 295-297). In this dissertation, we do not have the existence conditions for sequential games either. We may attempt to obtain solutions by taking limits on K_k^i defined in (2.7) and (2.8), provided the limits exist. It is possible that there are other solutions, which also provide Nash strategies.

Remark 4 *The condition in Theorem 3 is equivalent to that the associated coupled Riccati equations admit a positive semi-definite solution, which is known. Since in the infinite-horizon case ($N \rightarrow \infty$), we know that $x_N = 0$ due to the state feedback formulation. So we can choose arbitrary Q_N^1 and Q_N^2 . If the associated coupled Riccati equations admit a positive semi-definite solution (K_o^1, K_o^2) and (K_e^1, K_e^2) , we can choose $Q_N^i = K_o^i$ if N is odd-valued, or $Q_N^i = K_e^i$ if N is even-valued. Then the condition in Theorem 3 is satisfied. On the other hand, if the condition in Theorem 3 holds, (K_o^1, K_o^2) and (K_e^1, K_e^2) are the solutions to the associated coupled Riccati equations.*

2.2.2 Stability Analysis of State Feedback Nash Strategies

Theorem 4 *In Theorem 3, if K_e^1 , K_o^1 , K_e^2 , and K_o^2 are positive definite matrices, then the feedback system of the game defined in (2.50)-(2.53) with the Nash strategies (2.54)-(2.57) is uniformly asymptotically stable.*

Proof: Define the time varying Lyapunov function as

$$V_k(x_k) = x_k^\top (\bar{K}_k^1 + \bar{K}_k^2) x_k$$

where

$$\bar{K}_k^i = \begin{cases} K_e^i & , \text{ if } k \text{ is even} \\ K_o^i & , \text{ if } k \text{ is odd} \end{cases} \quad (2.58)$$

By the conditions $K_e^i > 0$ and $K_o^i > 0$, we obtain

$$W_1(x_k) \leq V_k(x_k) \leq W_2(x_k) \quad (2.59)$$

where $W_1(x_k) = \min(x_k^\top(K_e^1 + K_e^2)x_k, x_k^\top(K_o^1 + K_o^2)x_k)$, which is positive definite.

$W_2(x_k) = x_k^\top(K_e^1 + K_e^2 + K_o^1 + K_o^2)x_k$, which is also positive definite.

Next, we will prove $V_{k+1}(x_{k+1}) - V_k(x_k) \leq 0$. Let us consider the odd-valued k first.

$$\begin{aligned} V_{k+1}(x_{k+1}) - V_k(x_k) &= x_{k+1}^\top(K_e^1 + K_e^2)x_{k+1} - x_k^\top(K_o^1 + K_o^2)x_k \\ &= x_k^\top(A + B^1\bar{L}_k^1)^\top(K_e^1 + K_e^2)(A + B^1\bar{L}_k^1)x_k - x_k^\top(K_o^1 + K_o^2)x_k \\ &= -x_k^\top(Q^1 + \bar{L}_k^{1\top}R^{11}\bar{L}_k^1)x_k - x_k^\top(Q^2 + \bar{L}_k^{1\top}R^{21}\bar{L}_k^1)x_k \end{aligned}$$

Since $R^{11} > 0$, $R^{21} > 0$ and $Q^i \geq 0$, we can obtain $V_{k+1}(x_{k+1}) - V_k(x_k) \leq 0$, which implies the uniform stability of the zero equilibrium.

Finally, an additional argument is needed to establish asymptotic stability. Since (A, C^1) and (A, C^2) are observable (where $C^{1\top}C^1 = Q^1$ and $C^{2\top}C^2 = Q^2$), it is not possible that $y_k^1(x_k) = C^1x_k \equiv 0$ and $y_k^2(x_k) = C^2x_k \equiv 0$ unless $x_k \equiv 0$. Then

$$\begin{aligned} V_{k+1}(x_{k+1}) - V_k(x_k) &= -x_k^\top(Q^1 + \bar{L}_k^{1\top}R^{11}\bar{L}_k^1)x_k - x_k^\top(Q^2 + \bar{L}_k^{1\top}R^{21}\bar{L}_k^1)x_k \\ &\leq -(y_k^{1\top}y_k^1 + y_k^{2\top}y_k^2) < 0 \end{aligned}$$

We can obtain the same conclusion for even-valued k . So the system is uniformly asymptotically stable. ■

Remark 5 *Using a similar notation in Section 2.1.3, we can extend the results here to the M -player cases.*

2.3 Analysis of the Coupled Riccati Equations in the Scalar LQSG Cases

It is well known that there are no general results ensuring the existence of the solutions to the coupled Riccati equations in simultaneous closed loop Nash games. However, for a scalar case, we will show that a positive definite solution is guaranteed for the scalar LQSGs.

Theorem 5 *For the scalar case of Linear Quadratic Sequential Game defined as*

$$\begin{aligned} x_{k+1} &= ax_k + b^1 u_k^1 + b^2 u_k^2 \\ J^1 &= \sum_{i=0}^{\infty} (q^1 x_i^2 + r^{11} (u_i^1)^2 + r^{12} (u_i^2)^2) \\ J^2 &= \sum_{i=0}^{\infty} (q^2 x_i^2 + r^{21} (u_i^1)^2 + r^{22} (u_i^2)^2) \end{aligned}$$

where $q^i \geq 0$ and $r^{ij} > 0$ for $i = 1, 2$ and $j = 1, 2$, (a, b^1) and (a, b^2) are stabilizable, $(a, \sqrt{q^1})$ and $(a, \sqrt{q^2})$ are detectable, a positive solution to the related coupled Riccati Equations exists.

Proof: Since (a, b^1) and (a, b^2) are stabilizable, $b^1 \neq 0$ and $b^2 \neq 0$. W.L.O.G, we assume that $b^1 > 0$ and $b^2 > 0$. Otherwise, we can let $\bar{b}^i = -b^i$ and $\bar{u}_k^i = -u_k^i$. Then

the coupled discrete-time Algebraic Riccati Equations for the scalar cases are

$$K_e^1 = q^1 + (L_e^2)^2 r^{12} + (a + b^2 L_e^2)^2 K_o^1 \quad (2.60)$$

$$K_o^1 = q^1 + (L_o^1)^2 r^{11} + (a + b^1 L_o^1)^2 K_e^1 \quad (2.61)$$

$$K_e^2 = q^2 + (L_e^2)^2 r^{22} + (a + b^2 L_e^2)^2 K_o^2 \quad (2.62)$$

$$K_o^2 = q^2 + (L_o^1)^2 r^{21} + (a + b^1 L_o^1)^2 K_e^2 \quad (2.63)$$

Since

$$L_o^1 = \frac{-ab^1 K_e^1}{r^{11} + (b^1)^2 K_e^1}, \quad L_e^2 = \frac{-ab^2 K_o^2}{r^{22} + (b^2)^2 K_o^2},$$

we have

$$K_e^1 = q^1 + \frac{(b^2)^2 (K_o^2)^2 a^2 r^{12} + a^2 (r^{22})^2 K_o^1}{[r^{22} + (b^2)^2 K_o^2]^2} \quad (2.64)$$

$$K_o^1 = q^1 + \frac{(b^1)^2 (K_e^1)^2 a^2 r^{11} + a^2 (r^{11})^2 K_e^1}{[r^{11} + (b^1)^2 K_e^1]^2} \quad (2.65)$$

$$K_e^2 = q^2 + \frac{(b^2)^2 (K_o^2)^2 a^2 r^{22} + a^2 (r^{22})^2 K_o^2}{[r^{22} + (b^2)^2 K_o^2]^2} \quad (2.66)$$

$$K_o^2 = q^2 + \frac{(b^1)^2 (K_e^1)^2 a^2 r^{21} + a^2 (r^{11})^2 K_e^2}{[r^{11} + (b^1)^2 K_e^1]^2} \quad (2.67)$$

Denote $K_e := K_e^1$ and $K_o := K_o^2$. Simplifying (2.65) and (2.66), we get

$$K_o^1 = q^1 + \frac{a^2 K_e r^{11}}{r^{11} + (b^1)^2 K_e} \quad (2.68)$$

$$K_e^2 = q^2 + \frac{a^2 K_o r^{22}}{r^{22} + (b^2)^2 K_o} \quad (2.69)$$

By applying (2.68) and (2.69) to (2.64) and (2.67), respectively, we obtain

$$A_e(K_e)^2 - B_e K_e - C_e = 0 \quad (2.70)$$

$$A_o(K_o)^2 - B_o K_o - C_o = 0 \quad (2.71)$$

where,

$$A_e = (r^{22} + (b^2)^2 K_o)^2 b^1 \quad (2.72)$$

$$\begin{aligned} B_e &= (r^{22} + (b^2)^2 K_o)^2 (q^1 b^1 - r^{11}) + a^4 (r^{22})^2 r^{11} \\ &+ ((b^2)^2 (K_o)^2 r^{12} + (r^{22})^2 q^1) b^1 a^2 \end{aligned} \quad (2.73)$$

$$C_e = (r^{22} + (b^2)^2 K_o)^2 q^1 r^{11} + ((b^2)^2 (K_o)^2 r^{12} + (r^{22})^2 q^1) r^{11} a^2 \quad (2.74)$$

and

$$A_o = (r^{11} + (b^1)^2 K_e)^2 b^2 \quad (2.75)$$

$$\begin{aligned} B_o &= (r^{11} + (b^1)^2 K_e)^2 (q^2 b^2 - r^{22}) + a^4 (r^{11})^2 r^{22} \\ &+ ((b^1)^2 (K_e)^2 r^{21} + (r^{11})^2 q^2) b^2 a^2 \end{aligned} \quad (2.76)$$

$$C_o = (r^{11} + (b^1)^2 K_e)^2 q^2 r^{22} + ((b^1)^2 (K_e)^2 r^{21} + (r^{11})^2 q^2) r^{22} a^2 \quad (2.77)$$

Since $(a, \sqrt{q^1})$ and $(a, \sqrt{q^2})$ are detectable, $q^1 > 0$ and $q^2 > 0$. Then for any non-negative K_e and K_o , $A_e > 0$, $A_o > 0$, $C_e > 0$ and $C_o > 0$. For equations (2.70) and (2.71), we have

$$\Delta_e = B_e^2 + 4A_e C_e > 0 \quad (2.78)$$

$$\Delta_o = B_o^2 + 4A_o C_o > 0 \quad (2.79)$$

So, the roots of (2.70) and (2.71) are real-valued,

$$K_e = \frac{1}{2A_e} (B_e \pm \sqrt{\Delta_e}) \quad (2.80)$$

$$K_o = \frac{1}{2A_o} (B_o \pm \sqrt{\Delta_o}) \quad (2.81)$$

From (2.78) and (2.79), we know $\sqrt{\Delta_e} > |B_e|$ and $\sqrt{\Delta_o} > |B_o|$. So we can always find the positive valued roots of (2.70) and (2.71):

$$K_e = \frac{1}{2A_e}(B_e + \sqrt{\Delta_e}) := f_e(K_o) \quad (2.82)$$

$$K_o = \frac{1}{2A_o}(B_o + \sqrt{\Delta_o}) := f_o(K_e) \quad (2.83)$$

Remark 6 Only non-negative K_o and K_e are evaluated in function f_e and f_o , respectively.

Next, we will show that $\exists \bar{w}_e > 0$, such that $\bar{w}_e \geq K_e > 0$ for any $K_o \in [0, +\infty)$. Let $\bar{K}_o = r^{22} + (b^2)^2 K_o$, then we rewrite f_e as a function of $\bar{K}_o$, denote it $\bar{f}_e(\bar{K}_o)$.

$$\begin{aligned} K_e &= \frac{1}{2b^1(\bar{K}_o)^2} \left[\sum_{i=0}^2 \beta_i(\bar{K}_o)^i + \sqrt{\sum_{i=0}^4 \alpha_i(\bar{K}_o)^i} \right] \\ &= \frac{1}{2b^1} \left[\beta_2 + \frac{\beta_1}{\bar{K}_o} + \frac{\beta_0}{(\bar{K}_o)^2} + \sqrt{\sum_{i=0}^4 \frac{\alpha_i}{(\bar{K}_o)^{4-i}}} \right] \\ &\leq \frac{1}{2b^1} \left[\sum_{i=0}^2 \frac{|\beta_i|}{(\bar{K}_o)^{2-i}} + \sqrt{\sum_{i=0}^4 \frac{|\alpha_i|}{(\bar{K}_o)^{4-i}}} \right] \end{aligned} \quad (2.84)$$

where β_i and α_i are constants which depend on a , b^i , r^{ij} and q^i . Since $K_o \in [0, +\infty)$ and $r^{22} > 0$, $\bar{K}_o \geq r^{22} > 0$. Then

$$\bar{w}_e = \frac{1}{2b^1} \left[\sum_{i=0}^2 \frac{|\beta_i|}{(r^{22})^{2-i}} + \sqrt{\sum_{i=0}^4 \frac{|\alpha_i|}{(r^{22})^{4-i}}} \right] \geq K_e > 0 \quad (2.85)$$

So, function $f_e(K_o)$ is bounded from above by $\bar{w}_e$, and from below by 0. Similarly, we can find $\bar{w}_o > 0$, such that

$$\bar{w}_o \geq K_o > 0 \quad (2.86)$$

for any $K_e \in [0, +\infty)$.

From (2.85) and (2.86), we conclude that there is at least one positive pair (K_e, K_o)

which satisfies equations (2.70) and (2.71). We denote a pair as $(\mathcal{P}_e^1, \mathcal{P}_o^2)$. Then from (2.68) and (2.69),

$$\mathcal{P}_o^1 := K_o^1(\mathcal{P}_e^1) = q^1 + \frac{a^2 \mathcal{P}_e^1 r^{11}}{r^{11} + (b^1)^2 \mathcal{P}_e^1} \quad (2.87)$$

$$\mathcal{P}_e^2 := K_e^2(\mathcal{P}_o^2) = q^2 + \frac{a^2 \mathcal{P}_o^2 r^{22}}{r^{22} + (b^2)^2 \mathcal{P}_o^2} \quad (2.88)$$

Since $(a, \sqrt{q^1})$ and $(a, \sqrt{q^2})$ are detectable, $q^1 > 0$ and $q^2 > 0$. So, $\mathcal{P}_o^1 > 0$ and $\mathcal{P}_e^2 > 0$. Then,

$$\begin{cases} K_e^1 = \mathcal{P}_e^1 \\ K_o^1 = \mathcal{P}_o^1 \\ K_e^2 = \mathcal{P}_e^2 \\ K_o^2 = \mathcal{P}_o^2 \end{cases} \quad (2.89)$$

is a positive solution to the related coupled Riccati Equations, which completes our proof. ■

Remark 7 *From Theorem 4 and Theorem 5, we can always find state feedback Nash Strategies for scalar LQSGs such that the overall systems are uniformly asymptotically stable.*

2.4 Computer Simulations

Numerical examples are given in this section to illustrate our state feedback Nash strategies of LQSG switched systems.

2.4.1 2-order Systems

Consider the system defined in (2.50) - (2.53) with

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, B_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

and $R_{11} = R_{22} = 1, R_{12} = R_{21} = 2$. We can calculate K_k^1 and K_k^2 in Definition 1 and find that for any fixed odd-valued k , $K_k^1 \rightarrow K_o^1$ and $K_k^2 \rightarrow K_o^2$, and for any fixed odd-valued k , $K_k^1 \rightarrow K_o^1$ and $K_k^2 \rightarrow K_o^2$ as $N \rightarrow \infty$.

$$K_o^1 = \begin{bmatrix} 4.9829 & 5.2619 \\ 5.2619 & 8.1701 \end{bmatrix}, K_e^1 = \begin{bmatrix} 19.698 & 23.227 \\ 23.227 & 29.964 \end{bmatrix}$$

$$K_o^2 = \begin{bmatrix} 9.9557 & 10.516 \\ 10.516 & 17.416 \end{bmatrix}, K_e^2 = \begin{bmatrix} 29.392 & 32.671 \\ 32.671 & 42.058 \end{bmatrix}$$

The simulation results are shown in Fig. 2.1-Fig. 2.3 (in the following graphs, the plots are defined only at integer values of k). The feedback system with Nash Strategy is asymptotically stable.

Let us consider another numerical example,

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, B_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q_2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

and $R_{11} = R_{22} = 1, R_{12} = R_{21} = 2$. Then, we find that both K_k^1 and K_k^2 diverge. The feedback system is not stable as shown in Fig. 2.4 - Fig. 2.6. Actually, the system is not stabilizable.

2.4.2 A Scalar System

The following scalar numerical example will be used to illustrate the conclusion stated in *Remark 7*. The scalar system has $a = 0.3$, $b^1 = 2.3$, $b^2 = 0.5$, $q^1 = 1.5$, $q^2 = 0.6$, $r^{11} = 4.4$, $r^{12} = 4.7$, $r^{21} = 0.1$ and $r^{22} = 0.5$. As shown in Fig. 2.7, Fig. 2.8, and Fig. 2.9, at least one positive solution exist for the associated coupled Riccati equations. The closed-loop system with our state feedback Nash Strategies is stable, as shown in Fig. 2.10.

Figure 2.1: Asymptotically stable feedback system with Nash Strategy

Figure 2.2: K_k^1 , where $(K_k^1)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^1 , which is obtained from Definition 1. Positive definite K_e^1 and K_o^1 defined in Theorem 4 exist.

Figure 2.3: K_k^2 , where $(K_k^2)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^2 , which is obtained from Definition 1. Positive definite K_e^2 and K_o^2 defined in Theorem 4 exist.

Figure 2.4: Unstable feedback system with Nash Strategy

Figure 2.5: K_k^1 , where $(K_k^1)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^1 , which is obtained from Definition 1. K_k^1 diverges.

Figure 2.6: K_k^2 , where $(K_k^2)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^2 , which is obtained from Definition 1. K_k^2 diverges.

Figure 2.7: Function f_e and f_o

Figure 2.8: At least one solution exists. $K^2(1)$ and $K^1(2)$ are steady states of K_k^i shown in Fig. 2.9

Figure 2.9: K_k^1 and K_k^2 when $N = 11$

Figure 2.10: The feedback system is stable

2.5 Chapter Summary

In this section, a mathematical analysis of linear quadratic sequential games (LQSGs) was presented. A state feedback Nash strategy of the LQSGs was provided by using dynamic programming. A Hamiltonian based method was explored to derive the open-loop Nash strategies under perfect information structure. Finite and infinite time horizon cases were considered. For infinite-horizon sequentially switched systems, we proved uniform asymptotic stability of the state feedback Nash equilibrium using the Lyapunov's method, under the assumption that the DAREs have a positive definite solution. This is similar to the assumptions in the simultaneous game case. For scalar cases, we showed that a positive definite solution to the associated coupled DAREs is guaranteed for the stabilizable and detectable cases.

CHAPTER 3

ADAPTATION DESIGN FOR LINEAR QUADRATIC GAMES WITH STATE FEEDBACK NASH STRATEGIES

3.1 Introduction

Every closed-loop controller is “adaptive” in the sense that it modifies its output in response to changes of system states. However, an “adaptive controller” can adapt not only its output, but its underlying control strategy as well. Adaptive controllers can outperform their fixed-parameter counterparts in terms of efficiency. They can often eliminate errors faster and with fewer fluctuations. In [45], an adaptive controller is defined as “a controller with adjustable parameters and a mechanism for adjusting the parameters”. Usually, an adaptive control system has two loops: one is the information feedback between the process and the controller; the other is for the parameter adjustment.

In the control domain, where every system only has one controller, a range of techniques, such as adaptive dual control methods [49, 50], normalization or dead zones [51, 52], model reference adaptive control [53, 54, 55], and self-tuning regulators [56], are used to design the adaptive schemes. While in the game domain, there are two concepts related to the adaptation design. The first one is called evolutionary game theory (EGT) [57] that originated as an application of the mathematical theory

of games to biological contexts. Arising from the realization that frequency dependent fitness introduces a strategic aspect to evolution, EGT is the application of population genetics-inspired models of change in gene frequency in populations to game theory. It differs from classical game theory by focusing on the dynamics of strategy change more than the properties of strategy equilibria. The second one is Fictitious Play (FP) [6, 7]. In it, each player presumes that its opponents are playing stable (possibly mixed) strategies. Each player starts with some initial beliefs and chooses a best response to those beliefs as a strategy in this round. Then, after observing their opponents' actions, the players update their beliefs according to some learning rule (e.g. reinforcement learning or Bayes' rule). The process is then repeated.

In this chapter, we propose a new adaptive approach for linear quadratic games with state feedback Nash strategies. We consider the effort of adaptation under a FP framework with learning algorithms derived from conventional adaptive control methods. The remainder of the chapter is organized as follows. In **Section 3.2**, we consider a one-side adaptation scheme, in which only one player of the two players has perfect information of the cost functions of both players, while the other player only has access to its cost function and does not know the parameters of the cost function of its opponent. We design the adaptive law for the player without perfect information set in Section 3.2.1. Convergence and excitation conditions [45] are considered in Section 3.2.2. We extend the approach to LQSGs in Section 3.2.3. Computer simulations are provided in Section 3.2.4. In **Section 3.3**, we consider a two-side adaptation scheme, in which each player has only access to its own cost function as well as the system dynamics, and does not know the parameters of the cost function of its opponent, under a FP framework, in which we assume that each player

knows in advance that its opponent will implement Nash strategies. Adaptation design, convergence, extension to LQSGs, and computer simulations are provided in Section 3.3.1, Section 3.3.2, Section 3.3.3, and Section 3.3.4, respectively.

3.2 One-side Adaptation Scheme

Let us first consider two-person infinite-horizon simultaneous linear quadratic games

$$x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2 \quad (3.1)$$

with the cost functions

$$J^1 = \sum_{k=0}^{+\infty} (x_k^\top Q^1 x_k + u_k^{1\top} R^{11} u_k^1 + u_k^{2\top} R^{12} u_k^2) \quad (3.2)$$

$$J^2 = \sum_{k=0}^{+\infty} (x_k^\top Q^2 x_k + u_k^{1\top} R^{21} u_k^1 + u_k^{2\top} R^{22} u_k^2) \quad (3.3)$$

where $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}$, $Q^i > 0$, $R^{ij} > 0$, (A, B^i) is stabilizable, (A, C^i) is detectable (where $C^{i\top} C^i = Q^i$), and $|B^{i\top} B^i| > 0$ for $1 \leq i \leq 2$, $0 \leq k < +\infty$. We assume that both players have perfect information structures, with which both players know the exact system dynamics $x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2$ and measure exact system states x_k . It is also assumed that the simultaneous game in (3.1)-(3.3) has a unique Nash strategy pair for both players. It is well known (see [37]) that the Nash strategy pair is specified by

$$u_k^{1*} := \arg \min_{u_k^1} J^1 = \gamma^1(x_k) = L^1 x_k \quad (3.4)$$

$$u_k^{2*} := \arg \min_{u_k^2} J^2 = \gamma^2(x_k) = L^2 x_k \quad (3.5)$$

where L^1 and L^2 are defined, respectively, by

$$L^1 = -[R^{11} + B^{1\top} K^1 B^1]^{-1} B^{1\top} K^1 [A + B^2 L^2] \quad (3.6)$$

$$L^2 = -[R^{22} + B^{2\top} K^2 B^2]^{-1} B^{2\top} K^2 [A + B^1 L^1] \quad (3.7)$$

K^1 and K^2 are specified in the following coupled DAREs, that, by assumption, have a unique positive semi-definite solution.

$$K^1 = Q^1 + L^{1\top} R^{11} L^1 + L^{2\top} R^{12} L^2 + (A + B^1 L^1 + B^2 L^2)^\top K^1 (A + B^1 L^1 + B^2 L^2) \quad (3.8)$$

$$K^2 = Q^2 + L^{1\top} R^{21} L^1 + L^{2\top} R^{22} L^2 + (A + B^1 L^1 + B^2 L^2)^\top K^2 (A + B^1 L^1 + B^2 L^2) \quad (3.9)$$

In the one-side adaptation scheme, only one of the two players has perfect information of the cost functions of both players, i.e., the player knows exact Q^1 , Q^2 , R^{11} , R^{12} , R^{21} , R^{22} , while the other player only has access to its cost function and does not know the parameters of the cost function of its opponent. WLOG, we assume that player 1 knows exact Q^i , R^{ij} , A , B^i while player 2 has access to Q^2 , R^{21} , R^{22} , A , B^i . We also assume that player 1 will apply its state feedback Nash strategies calculated based the information of system dynamics, system states and cost functions of both players. Player 2 knows in advance that its opponent will implement state feedback Nash strategies. Player 2 will estimate the control gain L^1 first, and then calculate/estimate its own control gain L^2 .

3.2.1 Adaptation Design

In the one-side adaptation scheme, player 1, who has perfect information structure, will apply its real state feedback Nash strategies $u^{1*} = L^1 x_k$, so we follow the convectional indirect adaptive control design method [58].

Consider the system defined in (3.1) with fixed controller (3.4) for player 1, let $\hat{L}_k^1$ be an estimate of the control gain L^1 of player 1, from the point view of player 2. The block diagram of indirect adaptive control system is shown in Fig. 3.1.

First, we have

$$\hat{x}_{k+1} = Ax_k + B^1 \hat{L}_k^1 x_k + B^2 u_k^2 \quad (3.10)$$

Since matrix $B^{1\top} B^1$ is invertible, we obtain, from (3.1) and (3.10),

$$[B^{1\top} B^1]^{-1} B^{1\top} (\hat{x}_{k+1} - x_{k+1}) = (\hat{L}_k^1 - L^1) x_k \quad (3.11)$$

We introduce the estimation error $e_k = [B^{1\top} B^1]^{-1} B^{1\top} (\hat{x}_k - x_k)$, (e_k is a scalar) then

$$e_k = \tilde{L}_{k-1}^1 \phi_k \quad (3.12)$$

where $\tilde{L}_{k-1}^1 = \hat{L}_{k-1}^1 - L^1$ and $\phi_k = \frac{1}{z}[x]_k = x_{k-1}$. (Note, $\frac{1}{z}[x]_k$ denotes the output of the system with transfer function $\frac{1}{z}$ and input x_k .)

Figure 3.1: Indirect adaptive control design for one-side adaptation scheme

Choose the adaptive law for $\hat{L}_k^1$ as

$$\hat{L}_k^1 = \hat{L}_{k-1}^1 - \frac{e_k \phi_k^\top \Gamma}{m_k^2} \quad (3.13)$$

where $0 < \Gamma = \Gamma^\top < 2I_n$ is a gain matrix, and

$$m_k = \sqrt{\kappa + \phi_k^\top \phi_k}, \quad \kappa > 0$$

Convergence is proved and discussed in Section 3.2.2.

As shown in Fig. 3.1, for any $\hat{L}_k^1$, a design function or mapping will be used to calculate $\hat{L}_k^2$. From (3.7) and (3.9), we have a best response strategy for player 2 given the estimated control gain $\hat{L}_k^1$ of player 1,

$$\hat{L}_k^2 = -[R^{22} + B^{2\top} \hat{K}_k^2 B^2]^{-1} B^{2\top} \hat{K}_k^2 [A + B^1 \hat{L}_k^1] \quad (3.14)$$

$$\hat{K}_k^2 = Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1 + \hat{L}_k^{2\top} R^{22} \hat{L}_k^2 + (A + B^1 \hat{L}_k^1 + B^2 \hat{L}_k^2)^\top \hat{K}_k^2 (A + B^1 \hat{L}_k^1 + B^2 \hat{L}_k^2) \quad (3.15)$$

Lemma 3 *The $\hat{K}_k^2$ specified in (3.14) and (3.15) is the solution to the discrete-time algebraic Riccati equation (DARE)*

$$\bar{A}^\top X \bar{A} - X - \bar{A}^\top X \bar{B} [\bar{B}^\top X \bar{B} + \bar{R}]^{-1} \bar{B}^\top X \bar{A} + \bar{Q} = 0$$

where $\bar{A} = A + B^1 \hat{L}_k^1$, $\bar{B} = B^2$, $\bar{R} = R^{22}$, and $\bar{Q} = Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1$. And with the regularity condition specified below (3.3), i.e., $Q^i > 0$, $R^{ij} > 0$, (A, B^i) is stabilizable, (A, C^i) is detectable (where $C^{i\top} C^i = Q^i$), and $|B^{i\top} B^i| > 0$ (that is $B^{i\top} B^i$ is invertible), there exists a positive semi-definite solution to the above discrete time algebraic Riccati equation.

Proof: By matrix manipulation, we can rewrite (3.15) as

$$\begin{aligned} \hat{K}_k^2 &= \bar{Q} + \hat{L}_k^{2\top} \bar{R} \hat{L}_k^2 + (\bar{A} + \bar{B} \hat{L}_k^2)^\top \hat{K}_k^2 (\bar{A} + \bar{B} \hat{L}_k^2) \\ &= \bar{Q} + \bar{A}^\top \hat{K}_k^2 \bar{A} + \hat{L}_k^{2\top} [\bar{R} + \bar{B}^\top \hat{K}_k^2 \bar{B}] \hat{L}_k^2 + \bar{A}^\top \hat{K}_k^2 \bar{B} \hat{L}_k^2 + \hat{L}_k^{2\top} \bar{B}^\top \hat{K}_k^2 \bar{A} \\ &= \bar{Q} + \bar{A}^\top \hat{K}_k^2 \bar{A} + \bar{A}^\top \hat{K}_k^2 \bar{B} [\bar{R} + \bar{B}^\top \hat{K}_k^2 \bar{B}]^{-1} \bar{B}^\top \hat{K}_k^2 \bar{A} \\ &\quad - \bar{A}^\top \hat{K}_k^2 \bar{B} [\bar{R} + \bar{B}^\top \hat{K}_k^2 \bar{B}]^{-1} \bar{B}^\top \hat{K}_k^2 \bar{A} - \bar{A}^\top \hat{K}_k^2 \bar{B} [\bar{R} + \bar{B}^\top \hat{K}_k^2 \bar{B}]^{-1} \bar{B}^\top \hat{K}_k^2 \bar{A} \\ &= \bar{Q} + \bar{A}^\top \hat{K}_k^2 \bar{A} - \bar{A}^\top \hat{K}_k^2 \bar{B} [\bar{R} + \bar{B}^\top \hat{K}_k^2 \bar{B}]^{-1} \bar{B}^\top \hat{K}_k^2 \bar{A} \end{aligned} \quad (3.16)$$

In the above, the equality in the third line follows (3.14). Wonham [59] has proved that the existence condition of positive semi-definite solution to the discrete time

algebraic Riccati equation (3.16) is $(\bar{A}, \bar{B})$ is stabilizable and $(\bar{A}, \bar{C})$ is detectable (where $\bar{C}^\top \bar{C} = \bar{Q}$).

Since (A, B^2) is stabilizable, there exists a $\bar{L}$ such that $(A + B^2 \bar{L})$ is Hurwitz, that is it has all its eigenvalues inside the unit circle. Since $B^{2\top} B^2$ is invertible, there exists a $\bar{\bar{L}}$, such that

$$A + B^2 \bar{L} = \bar{A} + B^2 \bar{\bar{L}}$$

Actually, $\bar{\bar{L}} = (B^{2\top} B^2)^{-1} B^{2\top} (A + B^2 \bar{L} - \bar{A})$. Then $(\bar{A}, \bar{B})$ is stabilizable ($\bar{B} = B^2$).

Similarly, since (A, C^2) is detectable (where $C^{2\top} C^2 = Q^2$), we can always find a $\bar{p}$ such that $(A^\top + C^{2\top} \bar{p})$ is Hurwitz. From the condition $Q^2 > 0$, we have $\bar{Q} = Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1 > 0$. Then $\bar{C}^\top > 0$ where $\bar{C}^\top \bar{C} = \bar{Q}$. So there exist a $\bar{\bar{p}}$ such that

$$A^\top + C^{2\top} \bar{p} = \bar{A}^\top + \bar{C}^\top \bar{\bar{p}}$$

Actually $\bar{\bar{p}} = (\bar{C}^\top)^{-1} (A^\top + C^{2\top} \bar{p} - \bar{A}^\top)$. Then $(\bar{A}^\top + \bar{C}^\top \bar{\bar{p}})$ is Hurwitz, too. This completes our proof. ■

Remark 8 We can easily obtain $\hat{K}_k^2$ by using MATLAB command `dare` ($\bar{A}, \bar{B}, \bar{Q}, \bar{R}$).

3.2.2 Parameter Convergence and Persistency of Excitation

Similar to the proof of Normalized Gradient algorithm (see [58], page 115-116), we define a parameter error measurement function $V(\tilde{L}_k^1) = \tilde{L}_k^1 \Gamma^{-1} \tilde{L}_k^{1\top}$ for the adaptive law in (3.13). Then,

$$\begin{aligned} V(\tilde{L}_k^1) - V(\tilde{L}_{k-1}^1) &= \left[\tilde{L}_{k-1}^1 - \frac{e_k \phi_k^\top \Gamma}{m_k^2} \right] \Gamma^{-1} \left[\tilde{L}_{k-1}^1 - \frac{e_k \phi_k^\top \Gamma}{m_k^2} \right]^\top - \tilde{L}_{k-1}^1 \Gamma^{-1} \tilde{L}_{k-1}^{1\top} \\ &= -\frac{1}{m_k^2} \left(\tilde{L}_{k-1}^1 \phi_k e_k + e_k \phi_k^\top \tilde{L}_{k-1}^{1\top} - \frac{\phi_k^\top \Gamma \phi_k e_k^2}{m_k^2} \right) \\ &= -\frac{e_k^2}{m_k^2} \left(2 - \frac{\phi_k^\top \Gamma \phi_k}{\kappa + \phi_k^\top \phi_k} \right) \leq -\alpha \frac{e_k^2}{m_k^2} \end{aligned} \tag{3.17}$$

where $\alpha = 2 - \lambda_{\max}[\Gamma] > 0$. $\lambda_{\max}[\Gamma] \in (0, 2)$ is the maximum eigenvalue of Γ which satisfies the condition $0 < \Gamma = \Gamma^\top < 2I_n$. In the above, the equality in the first line comes from (3.13) and $\tilde{L}_k^1 = \hat{L}_k^1 - L^1$,

$$\hat{L}_k^1 = \hat{L}_{k-1}^1 - \frac{e_k \phi_k^\top \Gamma}{m_k^2} \Rightarrow \tilde{L}_k^1 = \tilde{L}_{k-1}^1 - \frac{e_k \phi_k^\top \Gamma}{m_k^2}$$

and the inequality in the last line can be testified by the following way. By the eigenvalue decomposition of Γ , we can obtain $\Gamma = W^\top \Lambda W$ where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is the eigenvalue matrix, and W contains the eigenvectors (W satisfies the condition $W^\top = W^{-1}$).

$$\phi_k^\top \Gamma \phi_k = \phi_k^\top W^\top \Lambda W \phi_k \leq \lambda_{\max}[\Gamma] \phi_k^\top W^\top W \phi_k = \lambda_{\max}[\Gamma] \phi_k^\top \phi_k$$

Since $\kappa > 0$, $\kappa + \phi_k^\top \phi_k \geq \phi_k^\top \phi_k$. Then

$$0 \leq \frac{\phi_k^\top \Gamma \phi_k}{\kappa + \phi_k^\top \phi_k} \leq \lambda_{\max}[\Gamma]$$

This completes the verification of (3.17), from which we have

$$\alpha \sum_{k=1}^N \frac{e_k^2}{m_k^2} \leq V(\hat{L}_0^1) - V(\hat{L}_N^1) \leq V(\hat{L}_0^1)$$

So $\lim_{N \rightarrow +\infty} \sum_{k=1}^N \frac{e_k^2}{m_k^2}$ is bounded from above by $V(\hat{L}_0^1)$. Since $m_k^2 = \kappa + \phi_k^\top \phi_k > 0$, we obtain

$$\lim_{k \rightarrow +\infty} e_k = \lim_{k \rightarrow +\infty} \tilde{L}_{k-1}^1 \phi_k = 0$$

If ϕ_k is persistently exciting, then

$$\lim_{k \rightarrow +\infty} \tilde{L}_k^1 = 0 \Rightarrow \lim_{k \rightarrow +\infty} \hat{L}_k^1 = L^1$$

We propose two ways to satisfy the excitation conditions (see [45], page 63-67). The first one is a reference signal tracking method, in which the cost functions of

decision makers need to be modified as follows:

$$J^1 = \sum_{k=0}^{+\infty} [(x_k - r_k)^\top Q^1 (x_k - r_k) + (u_k^1 - v_k^1)^\top R^{11} (u_k^1 - v_k^1) + (u_k^2 - v_k^2)^\top R^{12} (u_k^2 - v_k^2)] \quad (3.18)$$

$$J^2 = \sum_{k=0}^{+\infty} [(x_k - r_k)^\top Q^2 (x_k - r_k) + (u_k^1 - v_k^1)^\top R^{21} (u_k^1 - v_k^1) + (u_k^2 - v_k^2)^\top R^{22} (u_k^2 - v_k^2)] \quad (3.19)$$

To let the problem be well-defined (the optimal costs for both players are bounded), we require that $B^1 v_k^1 + B^2 v_k^2 = r_{k+1} - Ar_k$. (If the cost functions are bounded, then $k \rightarrow +\infty \Rightarrow x_k \rightarrow r_k$ and $u_k^i \rightarrow v_k^i \Rightarrow B^1 v_k^1 + B^2 v_k^2 = r_{k+1} - Ar_k$.)

Denoting $\bar{x}_k = x_k - r_k$ and $\bar{u}_k^i = u_k^i - v_k^i$, we can rewrite system state equation (3.1) and cost functions (3.18)-(3.19) as

$$\begin{aligned} \bar{x}_{k+1} &= A\bar{x}_k + B^1 \bar{u}_k^1 + B^2 \bar{u}_k^2 \\ J^1 &= \sum_{k=0}^{+\infty} (\bar{x}_k^\top Q^1 \bar{x}_k + \bar{u}_k^{1\top} R^{11} \bar{u}_k^1 + \bar{u}_k^{2\top} R^{12} \bar{u}_k^2) \\ J^2 &= \sum_{k=0}^{+\infty} (\bar{x}_k^\top Q^2 \bar{x}_k + \bar{u}_k^{1\top} R^{21} \bar{u}_k^1 + \bar{u}_k^{2\top} R^{22} \bar{u}_k^2) \end{aligned}$$

It is a standard simultaneous linear quadratic game problem with Nash strategies

$$\begin{aligned} \bar{u}_k^{1*} &:= \arg \min_{\bar{u}_k^1} J^1 = \gamma^1(\bar{x}_k) = L^1 \bar{x}_k \\ \bar{u}_k^{2*} &:= \arg \min_{\bar{u}_k^2} J^2 = \gamma^2(\bar{x}_k) = L^2 \bar{x}_k \end{aligned}$$

where L^1 and L^2 are defined in (3.6)-(3.7). Then

$$u_k^{i*} = L^i \bar{x}_k + v_k^i = L^i x_k + v_k^i - L^i r_k = L^i x_k + p_k^i$$

In this scenario, we assume that player 2 only needs to estimate the control gain L^1 .

Thus the excitation condition of the adaptive law in (3.13) is satisfied by choosing a proper reference input signal r_k .

The second method is to add small white noises w_k to the system state equation,

$$x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2 + w_k \quad (3.20)$$

then the associated cost functions are

$$J^1 = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=0}^N E(x_k^\top Q^1 x_k + u_k^{1\top} R^{11} u_k^1 + u_k^{2\top} R^{12} u_k^2) \quad (3.21)$$

$$J^2 = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{k=0}^N E(x_k^\top Q^2 x_k + u_k^{1\top} R^{21} u_k^1 + u_k^{2\top} R^{22} u_k^2) \quad (3.22)$$

The Nash strategies are the same as in (3.4)-(3.5). We choose the Recursive Least-Square estimator for player 2

$$\hat{L}_k^1 = \hat{L}_{k-1}^1 - e_k \phi_k^\top P_k \quad (3.23)$$

where $e_k = [B^{1\top} B^1]^{-1} B^{1\top} (\hat{x}_k - x_k) = \tilde{L}_{k-1}^1 \phi_k - [B^{1\top} B^1]^{-1} B^{1\top} w_{k-1}$ and

$$P_k = P_{k-1} - P_{k-1} \phi_k [I + \phi_k^\top P_{k-1} \phi_k]^{-1} \phi_k^\top P_{k-1}$$

with the sufficiently large initial condition $P_0 > 0$.

3.2.3 Extension to LQSGs

We can extend the adaptive methods in Section 3.2.1 and Section 3.2.2 to 2-player infinite-horizon LQSGs defined as

$$x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2 \quad (3.24)$$

with the constraints

$$\begin{cases} u_k^1 = 0, & \text{when } k \text{ is even} \\ u_k^2 = 0, & \text{when } k \text{ is odd} \end{cases} \quad (3.25)$$

and cost functions

$$J^1 = \sum_{i=0}^{+\infty} (x_i^\top Q^1 x_i + u_i^{1\top} R^{11} u_i^1 + u_i^{2\top} R^{12} u_i^2) \quad (3.26)$$

and

$$J^2 = \sum_{i=0}^{+\infty} (x_i^\top Q^2 x_i + u_i^{1\top} R^{21} u_i^1 + u_i^{2\top} R^{22} u_i^2) \quad (3.27)$$

where, $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}$, $Q^i \geq 0$, (A, C^i) are observable (where $C^{i\top} C^i = Q^i$), $R^{ij} > 0$ and $|B^{i\top} B^i| > 0$ for $i = 1, 2$ and $j = 1, 2$.

In the one-side adaptation scheme specified in the beginning part of Section 3.2, when k is odd, player 2 will update $\hat{L}_k^1$ according to the adaptive law (3.13) or (3.23), and when k is even, player 2 will calculate $\hat{L}_k^2$ based on $\hat{L}_k^1 = \hat{L}_{k-1}^1$. Hence we arrive at the following theorem.

Theorem 5 *Consider the one-side adaptation scheme of LQSGs defined in (3.24)-(3.27),*

(i) *for the reference input tracking case, the convergent adaptive law for player 2 is*

$$\hat{L}_k^1 = \hat{L}_{k-1}^1 - \frac{e_k \phi_k^{1\top} \Gamma}{m_k^2} \quad (3.28)$$

where $0 < \Gamma = \Gamma^\top < 2I_n$ is a gain matrix, and

$$m_k = \sqrt{\kappa + \phi_k^{1\top} \phi_k^1}, \quad \kappa > 0$$

$$\phi_k^1 = \begin{cases} 0 & , \text{ when } k \text{ is even} \\ x_{k-1} & , \text{ when } k \text{ is odd} \end{cases} \quad (3.29)$$

(ii) *for the second case, in which the system state equation has the small white noise as model error, the convergent adaptive law for player 2 is*

$$\hat{L}_k^1 = \hat{L}_{k-1}^1 - e_k \phi_k^{1\top} P_k \quad (3.30)$$

where $e_k = [B^{1\top} B^1]^{-1} B^{1\top} (\hat{x}_k - x_k)$ and

$$P_k = P_{k-1} - P_{k-1} \phi_k^1 [I + \phi_k^{1\top} P_{k-1} \phi_k^1]^{-1} \phi_k^{1\top} P_{k-1}$$

with the sufficiently large initial condition $P_0 > 0$ and ϕ_k^1 is defined in (3.29).

(iii) after getting $\hat{L}_k^1$ in case (i) or (ii), we can, for even-valued k , calculate $\hat{L}_k^2$ from the equation

$$\hat{L}_k^2 = -[R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \quad (3.31)$$

where $\hat{K}_{o,k}^2$, the estimates of K_o^2 defined in the Remark 3 of Section 2.2 at time k , is the positive semi-definite solution to the DARE,

$$\bar{A}^\top X \bar{A} - X - \bar{A}^\top X \bar{B} [\bar{B}^\top X \bar{B} + \bar{R}]^{-1} \bar{B}^\top X \bar{A} + \bar{Q} = 0$$

with $\bar{Q} = [A + B^1 \hat{L}_k^1]^\top Q^2 [A + B^1 \hat{L}_k^1] + Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1$, $\bar{A} = A[A + B^1 \hat{L}_k^1]$, $\bar{B} = B^2$, and $\bar{R} = R^{22}$.

Proof: The proof of (i) and (ii) follows the Section 3.2.1, Section 3.2.2 and the discussion in Section 3.2.3 (just before this Theorem). We only need to prove the conclusion in part (iii). From Remark 3 of Section 2.2, we recall that

$$\begin{aligned} L^2 &= -[R^{22} + B^{2\top} K_o^2 B^2]^{-1} B^{2\top} K_o^2 A \\ K_o^2 &= Q^2 + L^{1\top} R^{21} L^1 + (A + B^1 L^1)^\top K_e^2 (A + B^1 L^1) \\ K_e^2 &= Q^2 + L^{2\top} R^{22} L^2 + (A + B^2 L^2)^\top K_o^2 (A + B^2 L^2) \end{aligned}$$

Then, given $\hat{L}_k^1$, we can calculate $\hat{L}_k^2$ by

$$\begin{aligned} \hat{L}_k^2 &= -[R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \\ \hat{K}_{o,k}^2 &= Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1 + (A + B^1 \hat{L}_k^1)^\top \hat{K}_{e,k}^2 (A + B^1 \hat{L}_k^1) \\ \hat{K}_{e,k}^2 &= Q^2 + \hat{L}_k^{2\top} R^{22} \hat{L}_k^2 + (A + B^2 \hat{L}_k^2)^\top \hat{K}_{o,k}^2 (A + B^2 \hat{L}_k^2) \end{aligned}$$

So,

$$\begin{aligned}
\hat{K}_{e,k}^2 &= Q^2 + \hat{L}_k^{2\top} [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2] \hat{L}_k^2 + A^\top \hat{K}_{o,k}^2 A + \hat{L}_k^{2\top} B^{2\top} \hat{K}_{o,k}^2 A + A^\top \hat{K}_{o,k}^2 B^2 \hat{L}_k^2 \\
&= Q^2 + A^\top \hat{K}_{o,k}^2 A + A^\top \hat{K}_{o,k}^2 B^2 [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \\
&\quad - A^\top \hat{K}_{o,k}^2 B^2 [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \\
&\quad - A^\top \hat{K}_{o,k}^2 B^2 [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \\
&= Q^2 + A^\top \hat{K}_{o,k}^2 A - A^\top \hat{K}_{o,k}^2 B^2 [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \tag{3.32}
\end{aligned}$$

Multiplying $(A + B^1 \hat{L}_k^1)^\top$ and $(A + B^1 \hat{L}_k^1)$ to left and right side of (3.32), respectively, we obtain

$$\begin{aligned}
(A + B^1 \hat{L}_k^1)^\top \hat{K}_{e,k}^2 (A + B^1 \hat{L}_k^1) &= (A + B^1 \hat{L}_k^1)^\top Q^2 (A + B^1 \hat{L}_k^1) + \bar{A}^\top \hat{K}_{o,k}^2 \bar{A} \\
&\quad - \bar{A}^\top \hat{K}_{o,k}^2 B^2 [R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 \bar{A}
\end{aligned}$$

Adding $Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1$ to both sides of the previous equation, we have

$$\hat{K}_{o,k}^2 = \bar{Q} + \bar{A}^\top \hat{K}_{o,k}^2 \bar{A} - \bar{A}^\top \hat{K}_{o,k}^2 \bar{B} [\bar{R} + \bar{B}^\top \hat{K}_{o,k}^2 \bar{B}]^{-1} \bar{B}^\top \hat{K}_{o,k}^2 \bar{A} \tag{3.33}$$

Since $\bar{Q} = [A + B^1 \hat{L}_k^1]^\top Q^2 [A + B^1 \hat{L}_k^1] + Q^2 + \hat{L}_k^{1\top} R^{21} \hat{L}_k^1 > 0$, and $(\bar{B}^\top \bar{B})$ is invertible, we can prove that $(\bar{A}, \bar{B})$ is stabilizable and $(\bar{A}, \bar{C})$ is detectable (where $\bar{C}^\top \bar{C} = \bar{Q}$) by using the same method in the proof of Lemma 3. So the Riccati equation in (3.33) exists a positive semi-definite solution $\hat{K}_{o,k}^2$ by using the result in [59], in which Wonham has proved that the existence condition of a positive semi-definite solution to the discrete time algebraic Riccati equation (3.33) is: $(\bar{A}, \bar{B})$ is stabilizable and $(\bar{A}, \bar{C})$ is detectable (where $\bar{C}^\top \bar{C} = \bar{Q}$). This completes our proof of part (iii). $\blacksquare$

3.2.4 Simulations and Discussions

Numerical examples are given in this section to illustrate our one-side adaptation for the state feedback Nash strategies of infinite-horizon linear quadratic games.

Linear Quadratic Simultaneous Game Case

Consider the system defined in (3.1) - (3.3) with

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, Q^1 = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}, Q^2 = \begin{bmatrix} 1.6 & 0 \\ 0 & 2.1 \end{bmatrix}$$

$$B^1 = [1.3, 1.7]^\top, B^2 = [1.5, 1.6]^\top, R^{11} = 0.4, R^{12} = 1.0, R^{21} = 1.0, \text{ and } R^{22} = 0.5.$$

We can calculate K^1 , K^2 , L^1 and L^2 from the coupled discrete-time algebraic Riccati equation specified in (3.6)-(3.9),

$$K^1 = \begin{bmatrix} 2.3609 & 1.0957 \\ 1.0957 & 3.4812 \end{bmatrix}, K^2 = \begin{bmatrix} 2.5639 & 1.2059 \\ 1.2059 & 4.2965 \end{bmatrix}$$

$L^1 = [-0.4992, -1.1682]$ and $L^2 = [-0.5215, -1.1295]$. We first carry out the adaptation based on the adaptive law (3.13) for the system with cost functions in (3.18)-(3.19), where $\Gamma = I$, $\kappa = 0.4$ and

$$r_k = \begin{bmatrix} 1.7 \sin(0.5k) \\ \cos(0.9k) \end{bmatrix}$$

As shown in Fig. 3.2, we can see that $\hat{L}_k^1$, $\hat{L}_k^2$, and $\hat{x}_k$ converge to L^1 , L^2 and x_k , respectively.

Secondly, we implement the adaptive law in (3.23) with white noise

$$w \sim \mathcal{N}(\mu = [0, 0]^\top, \sigma^2 = 0.01I_2)$$

as the system disturbance. From Fig. 3.3, we can see unbiased estimates for L^1 and L^2 .

Figure 3.2: $\hat{L}^i$, L^i , x_k and $\hat{x}_k$ in one-side adaptation for a simultaneous game: reference signal tracking case. $\{\hat{L}_k^i\}_i$ is the i^{th} element of $\hat{L}_k^i$. The reference signal is $r_k = [1.7 \sin(0.5k), \cos(0.9k)]^\top$. We can see the convergence of $\hat{L}_k^1$ and $\hat{L}_k^2$.

Figure 3.3: $\hat{L}^i$, L^i , x_k and $\hat{x}_k$ in one-side adaptation a simultaneous game with small system noise. $\{\hat{L}_k^i\}_i$ is the i^{th} element of $\hat{L}_k^i$. The system noise is $w = \mathcal{N}([0, 0]^\top, 0.01I)$. The initial value $P_0 = 100I$. We can see the convergence of $E(\hat{L}_k^1)$ and $E(\hat{L}_k^2)$ from the average values of 20 runs.

Linear Quadratic Sequential Game Case

Let us consider the LQSG defined in (3.24)-(3.27) with the same A , B^i , Q^i , and R^{ij} as in the Linear Quadratic Simultaneous Game Case. We can calculate K_o^i , K_e^i , L^1 and L^2 from the coupled discrete-time algebraic Riccati equation specified in Remark 3 of Section 2.2,

$$K_o^1 = \begin{bmatrix} 2.2454 & 1.0086 \\ 1.0086 & 3.6708 \end{bmatrix}, K_e^1 = \begin{bmatrix} 3.3203 & 3.0207 \\ 3.0207 & 7.3355 \end{bmatrix}$$

$$K_o^2 = \begin{bmatrix} 3.0301 & 2.4535 \\ 2.4535 & 7.4830 \end{bmatrix}, K_e^2 = \begin{bmatrix} 3.0125 & 1.9535 \\ 1.9535 & 5.4182 \end{bmatrix}$$

$L^1 = [-1.0415, -2.3160]$ and $L^2 = [-1.0399, -2.3012]$. We first implement the adaptive scheme based on the adaptive law (3.28)-(3.29) with $\Gamma = I$, $\kappa = 0.3$ and

$$r_k = \begin{bmatrix} 1.7 \sin(0.5k) \\ \cos(0.9k) \end{bmatrix}$$

As shown in Fig. 3.4, we can see that $\hat{L}_k^1$, $\hat{L}_k^2$, and $\hat{x}_k$ converge to L^1 , L^1 and x_k , respectively.

Secondly, we implement the adaptive law in (3.30) with white noise $w \sim \mathcal{N}(\mu = [0, 0]^\top, \sigma^2 = 0.01I_2)$ as the system disturbance. From Fig. 3.5, we can see unbiased estimates for L^1 and L^2 based on the average values of 20 runs.

Figure 3.4: $\hat{L}^i$, L^i , x_k and $\hat{x}_k$ in one-side adaptation for LQSG: reference signal tracking case. $\{\hat{L}_k^i\}_i$ is the i^{th} element of $\hat{L}_k^i$. The reference signal is $r_k = [1.7 \sin(0.5k), \cos(0.9k)]^\top$. We can see the convergence of $\hat{L}_k^1$ and $\hat{L}_k^2$.

Figure 3.5: $\hat{L}^i$, L^i , x_k and $\hat{x}_k$ in one-side adaptation for a LQSG with small system noise. $\{\hat{L}_k^i\}_i$ is the i^{th} element of $\hat{L}_k^i$. The system noise is $w = \mathcal{N}([0, 0]^\top, 0.01I)$. The initial value $P_0 = 100I$. We can see the convergence of $E(\hat{L}_k^1)$ and $E(\hat{L}_k^2)$ from the average values of 20 runs.

3.3 Two-side Adaptation Scheme

For the same linear quadratic infinite horizon simultaneous game defined in (3.1)-(3.9), we provide adaptation design, parameter convergence, and persistence of excitation for the two-side adaptation scheme, in which each player needs to estimate the control gain of its opponent. We assume that each player has access to its own cost function as well as the system dynamics, and does not know the parameters of the cost function of its opponent. i.e., player 1 knows Q^1 , R^{12} , R^{11} , A , B^1 and B^2 while player 2 has access to Q^2 , R^{21} , R^{22} , A , B^1 and B^2 . We also assume that each player knows in advance that its opponent will implement state feedback Nash strategies. So, player 2 (or player 1) will estimate the control gain L^1 (or L^2) first, and then calculate/estimate its own control gain L^2 (or L^1).

3.3.1 Adaptation Design

Consider the system defined in (3.1), let $\hat{L}_k^{12}$ and $\hat{L}_k^{21}$ be, respectively, the estimate of the control gain L^2 and L^1 from the point view of player 1 and player 2. Player 1 then calculates the associated estimate $\hat{L}_k^{11}$ of its own control gain L^1 from $\hat{L}_k^{12}$. Similarly, player 2 calculates the associated estimate $\hat{L}_k^{22}$ of its own control gain L^2 from $\hat{L}_k^{21}$. The block diagram of indirect adaptive control system is shown in Fig. 3.6.

First we have

$$\hat{x}_{k+1}^1 = Ax_k + B^1 \hat{L}_k^{11} x_k + B^2 \hat{L}_k^{12} x_k \quad (3.34)$$

$$\hat{x}_{k+1}^2 = Ax_k + B^1 \hat{L}_k^{21} x_k + B^2 \hat{L}_k^{22} x_k \quad (3.35)$$

$$x_{k+1} = Ax_k + B^1 \hat{L}_k^{11} x_k + B^2 \hat{L}_k^{22} x_k \quad (3.36)$$

where x_k is the real system state, $\hat{x}_k^1$ and $\hat{x}_k^2$ are the estimates of system state for player 1 and player 2, respectively. Since matrix $B^{1\top}B^1$ and $B^{2\top}B^2$ are invertible, we obtain, from (3.34)-(3.36),

$$[B^{2\top}B^2]^{-1}B^{2\top}(\hat{x}_{k+1}^1 - x_{k+1}) = (\hat{L}_k^{12} - \hat{L}_k^{22})x_k \quad (3.37)$$

$$[B^{1\top}B^1]^{-1}B^{1\top}(\hat{x}_{k+1}^2 - x_{k+1}) = (\hat{L}_k^{21} - \hat{L}_k^{11})x_k \quad (3.38)$$

Figure 3.6: Indirect adaptive control design for two-side adaptation scheme

We introduce the estimation errors

$$e_k^1 = [B^{2\top}B^2]^{-1}B^{2\top}(\hat{x}_k^1 - x_k), \text{ and } e_k^2 = [B^{1\top}B^1]^{-1}B^{1\top}(\hat{x}_k^2 - x_k)$$

(e_k^i is a scalar) for player 1 and player 2, respectively, then

$$e_k^1 = \tilde{L}_{k-1}^2 \phi_k \quad (3.39)$$

$$e_k^2 = \tilde{L}_{k-1}^1 \phi_k \quad (3.40)$$

where $\tilde{L}_{k-1}^2 = \hat{L}_{k-1}^{12} - \hat{L}_{k-1}^{22}$, $\tilde{L}_{k-1}^1 = \hat{L}_{k-1}^{21} - \hat{L}_{k-1}^{11}$ and $\phi_k = \frac{1}{z}[x]_k = x_{k-1}$. (Note, $\frac{1}{z}[x]_k$ denotes the output of the system with transfer function $\frac{1}{z}$ and input x_k .)

Choose the adaptive law for $\hat{L}_k^{12}$ and $\hat{L}_k^{21}$ as

$$\hat{L}_k^{12} = \hat{L}_{k-1}^{12} - \frac{e_k^1 \phi_k^\top}{(m_k^1)^2} \Gamma_1 \quad (3.41)$$

$$\hat{L}_k^{21} = \hat{L}_{k-1}^{21} - \frac{e_k^2 \phi_k^\top}{(m_k^2)^2} \Gamma_2 \quad (3.42)$$

where $m_k^1 = \sqrt{\kappa_1 + \phi_k^\top \phi_k}$, $m_k^2 = \sqrt{\kappa_2 + \phi_k^\top \phi_k}$, $\kappa_1 > 0$, $\kappa_2 > 0$, $0 < \Gamma_1 = \Gamma_1^\top \leq I$, and $0 < \Gamma_2 = \Gamma_2^\top \leq I$. Convergence is proved and discussed in Section 3.3.2.

As shown in Fig. 3.6, for any $\hat{L}_k^{12}$ (or $\hat{L}_k^{21}$), a design function or mapping will be used to calculate the associated $\hat{L}_k^{11}$ (or $\hat{L}_k^{22}$). From (3.6)-(3.9), we have a best response strategy for player 1 (or player 2) given the estimated control gain $\hat{L}_k^{12}$ or $(\hat{L}_k^{21})$ of player 2 (or player 1) at time k ,

$$\hat{L}_k^{11} = -[R^{11} + B^{1\top} \hat{K}_k^1 B^1]^{-1} B^{1\top} \hat{K}_k^1 [A + B^2 \hat{L}_k^{12}] \quad (3.43)$$

$$\hat{L}_k^{22} = -[R^{22} + B^{2\top} \hat{K}_k^2 B^2]^{-1} B^{2\top} \hat{K}_k^2 [A + B^1 \hat{L}_k^{21}] \quad (3.44)$$

where

$$\hat{K}_k^1 = Q^1 + \hat{L}_k^{11\top} R^{11} \hat{L}_k^{11} + \hat{L}_k^{12\top} R^{12} \hat{L}_k^{12} + (A + B^1 \hat{L}_k^{11} + B^2 \hat{L}_k^{12})^\top \hat{K}_k^1 (A + B^1 \hat{L}_k^{11} + B^2 \hat{L}_k^{12})$$

$$\hat{K}_k^2 = Q^2 + \hat{L}_k^{21\top} R^{21} \hat{L}_k^{21} + \hat{L}_k^{22\top} R^{22} \hat{L}_k^{22} + (A + B^1 \hat{L}_k^{21} + B^2 \hat{L}_k^{22})^\top \hat{K}_k^2 (A + B^1 \hat{L}_k^{21} + B^2 \hat{L}_k^{22})$$

By Lemma 3, we can prove that $\hat{K}^1$ and $\hat{K}^2$ satisfy the following discrete time algebraic Riccati equations, respectively,

$$\bar{A}_1^\top \hat{K}_k^1 \bar{A}_1 - \hat{K}_k^1 - \bar{A}_1^\top \hat{K}_k^1 B^1 [B^{1\top} \hat{K}_k^1 B^1 + R^{11}]^{-1} B^{1\top} \hat{K}_k^1 \bar{A}_1 + \bar{Q}^1 = 0 \quad (3.45)$$

$$\bar{A}_2^\top \hat{K}_k^2 \bar{A}_2 - \hat{K}_k^2 - \bar{A}_2^\top \hat{K}_k^2 B^2 [B^{2\top} \hat{K}_k^2 B^2 + R^{22}]^{-1} B^{2\top} \hat{K}_k^2 \bar{A}_2 + \bar{Q}^2 = 0 \quad (3.46)$$

where $\bar{A}_1 = A + B^2 \hat{L}_k^{12}$, $\bar{Q}^1 = Q^1 + \hat{L}_k^{12\top} R^{12} \hat{L}_k^{12}$, $\bar{A}_2 = A + B^1 \hat{L}_k^{21}$, and $\bar{Q}^2 = Q^2 + \hat{L}_k^{21\top} R^{21} \hat{L}_k^{21}$.

Remark 9 We may obtain $\hat{K}_k^1$ and $\hat{K}_k^2$ by MATLAB command `dare`($\bar{A}_1, B^1, \bar{Q}^1, R^{11}$) and `dare`($\bar{A}_2, B^2, \bar{Q}^2, R^{22}$), respectively.

3.3.2 Parameter Convergence and Persistency of Excitation

To satisfy the excitation conditions (see [45], page 63-67), we propose a reference signal tracking method, in which the cost functions of decision makers need to be modified as specified in (3.18)-(3.19). As mentioned in the one-side adaptation scheme (see Section 3.2.2), we also require that $B^1 v_k^1 + B^2 v_k^2 = r_{k+1} - A r_k$. (If the cost functions are bounded, then $k \rightarrow +\infty \Rightarrow x_k \rightarrow r_k$ and $u_k^i \rightarrow v_k^i \Rightarrow B^1 v_k^1 + B^2 v_k^2 = r_{k+1} - A r_k$.)

It is already verified in Section 3.2.2 that under the perfect information structure the Nash strategies are

$$\begin{aligned} \bar{u}_k^{1*} &:= \arg \min_{\bar{u}_k^1} J^1 = \gamma^1(\bar{x}_k) = L^1 \bar{x}_k = L^1 x_k + p_k^1 \\ \bar{u}_k^{2*} &:= \arg \min_{\bar{u}_k^2} J^2 = \gamma^2(\bar{x}_k) = L^2 \bar{x}_k = L^2 x_k + p_k^2 \end{aligned}$$

where L^1 and L^2 are defined in (3.6)-(3.7).

In this scenario, we assume that each player only needs to estimate the control gain of its opponent. Thus the excitation conditions of the adaptive laws in (3.41)-(3.42) are satisfied by choosing a proper reference input signal r_k .

From Lemma 3, we know that for any $\hat{L}_k^{12}$ (or $\hat{L}_k^{21}$), we can obtain the unique positive semi-definite solution $\hat{K}_k^1$ (or $\hat{K}_k^2$), then we can calculate $\hat{L}_k^{11}$ (or $\hat{L}_k^{22}$) by using

equation (3.43) or (3.44). Let us define two functions as follows:

$$f^1 : \hat{L}_k^{12} \mapsto \hat{L}_k^{11} \quad (3.47)$$

$$f^2 : \hat{L}_k^{21} \mapsto \hat{L}_k^{22} \quad (3.48)$$

For an example, for the scalar case, the reaction-curve-like relation between f^1 and f^2 is shown in Fig. 3.7.

Definition 2 *Similar to the “reaction curve” concept of game theory, we define two properties as*

$$\text{Property I: } \begin{cases} \|f^1(f^2(a)) - L^1\|_2 & \leq \|a - L^1\|_2 \\ \|f^2(f^1(b)) - L^2\|_2 & \leq \|b - L^2\|_2 \end{cases}$$

$$\text{Property II: } \begin{cases} \|a_1 - L^1\|_2 \leq \|a_2 - L^1\|_2 & \Rightarrow \|f^2(a_1) - L^2\|_2 \leq \|f^2(a_2) - L^2\|_2 \\ \|b_1 - L^2\|_2 \leq \|b_2 - L^2\|_2 & \Rightarrow \|f^1(b_1) - L^1\|_2 \leq \|f^1(b_2) - L^1\|_2 \end{cases}$$

where $\|a\|_2$ is the standard Euclidean norm. For row vector a , $\|a\|_2 = aa^\top$. The equalities in Property I hold only when $a = L^1$ and $b = L^2$.

Remark 10 *Property I means that each player will use its reaction chance to improve its strategy in the sense that the distance from the current control gain to the optimal control gain is reduced at each reaction time.*

Remark 11 *Property II means that the distance (from the current control gain to the optimal control gain) relation between the two players is positive.*

Theorem 6 *For the system defined in (3.1) with objective functions in (3.18)-(3.19), if the associated coupled discrete time algebraic Riccati equation specified in (3.6)-(3.9) has a unique positive semi-definite solution, then the two-side adaptation with*

Figure 3.7: Reaction-curve-like relation between function $\hat{L}_k^{11} = f^1(\hat{L}_k^{12})$ and $\hat{L}_k^{22} = f^2(\hat{L}_k^{21})$. Since we assume that there exist a unique positive semi-definite solution to the coupled discrete time algebraic Riccati equations in (3.6)-(3.9), there is only one cross point (marked by a green cross in this plot), which is composed of the real control gains L^2 and L^1 of Player 2 and Player 1, respectively.

adaptive laws in (3.41)-(3.42) and design equations in (3.43)-(3.44) will guarantee convergence. i.e.,

$$\lim_{k \rightarrow +\infty} \hat{L}_k^{ij} = L^j$$

where $i = 1, 2$ and $j = 1, 2$.

Proof: We first prove the following two inequalities for the adaptive laws specified in (3.41)-(3.42),

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \max \left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2 \right) \quad (3.49)$$

$$\|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \max \left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2 \right) \quad (3.50)$$

From (3.41), we can obtain

$$\hat{L}_{k+1}^{12} - \hat{L}_k^{12} = -\frac{e_{k+1}^1 \phi_{k+1}^\top \Gamma_1}{(m_{k+1}^1)^2}$$

and

$$\begin{aligned} \hat{L}_{k+1}^{12} - \hat{L}_k^{22} &= [\hat{L}_k^{12} - \hat{L}_k^{22}] - \frac{e_{k+1}^1 \phi_{k+1}^\top \Gamma_1}{(m_{k+1}^1)^2} = [\hat{L}_k^{12} - \hat{L}_k^{22}] - \frac{[\hat{L}_k^{12} - \hat{L}_k^{22}] \phi_{k+1} \phi_{k+1}^\top \Gamma_1}{(m_{k+1}^1)^2} \\ &= [\hat{L}_k^{12} - \hat{L}_k^{22}] \left[I - \frac{\phi_{k+1} \phi_{k+1}^\top \Gamma_1}{(m_{k+1}^1)^2} \right] \end{aligned}$$

Then

$$\begin{aligned} (\hat{L}_{k+1}^{12} - \hat{L}_k^{12}) \Gamma_1^{-1} (\hat{L}_{k+1}^{12} - \hat{L}_k^{22})^\top &= -\frac{e_{k+1}^1 \phi_{k+1}^\top [\hat{L}_k^{12} - \hat{L}_k^{22}]^\top}{(m_{k+1}^1)^2} \\ &\quad + \frac{e_{k+1}^1 \phi_{k+1}^\top \Gamma_1 \phi_{k+1} \phi_{k+1}^\top [\hat{L}_k^{12} - \hat{L}_k^{22}]^\top}{(m_{k+1}^1)^4} \\ &= -\frac{(e_{k+1}^1)^2}{(m_{k+1}^1)^2} + \frac{(e_{k+1}^1)^2 \phi_{k+1}^\top \Gamma_1 \phi_{k+1}}{(m_{k+1}^1)^4} \\ &= -\frac{(e_{k+1}^1)^2}{(m_{k+1}^1)^2} \left[1 - \frac{\phi_{k+1}^\top \Gamma_1 \phi_{k+1}}{(m_{k+1}^1)^2} \right] \end{aligned}$$

Since $\phi_{k+1}^\top \Gamma_1 \phi_{k+1} \leq \lambda_{\max}[\Gamma_1] \phi_{k+1}^\top \phi_{k+1} \leq \phi_{k+1}^\top \phi_{k+1}$,

$$(\hat{L}_{k+1}^{12} - \hat{L}_k^{12}) \Gamma_1^{-1} (\hat{L}_{k+1}^{12} - \hat{L}_k^{22})^\top \leq -\frac{(e_{k+1}^1)^2}{(m_{k+1}^1)^2} \frac{\kappa_1}{(m_{k+1}^1)^2} \leq 0$$

So, $(\hat{L}_{k+1}^{12}, f^1(\hat{L}_{k+1}^{12}))$ is between $(\hat{L}_k^{12}, f^1(\hat{L}_k^{12}))$ and $(\hat{L}_k^{22}, f^1(\hat{L}_k^{22}))$. Therefore, (3.49)

is true. Similarly, we can prove (3.50) as well.

Next, we will prove the following two inequalities:

$$\max \left(\|\hat{L}_{k+1}^{12} - L^2\|_2, \|\hat{L}_{k+1}^{22} - L^2\|_2 \right) \leq \max \left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2 \right) \quad (3.51)$$

$$\max \left(\|\hat{L}_{k+1}^{21} - L^1\|_2, \|\hat{L}_{k+1}^{11} - L^1\|_2 \right) \leq \max \left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2 \right) \quad (3.52)$$

At time k , as shown in Fig. 3.8 for a scalar example, there are only three possible cases of the relation between $(\hat{L}_k^{12}, \hat{L}_k^{11})$ and $(\hat{L}_k^{22}, \hat{L}_k^{21})$:

Case 1: $\|\hat{L}_k^{11} - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2$ and $\|\hat{L}_k^{22} - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2$.

Case 2: $\|\hat{L}_k^{12} - L^2\|_2 < \|\hat{L}_k^{22} - L^2\|_2$.

Case 3: $\|\hat{L}_k^{21} - L^1\|_2 < \|\hat{L}_k^{11} - L^1\|_2$.

If Case 1 is true, then from (3.49) and (3.50), we obtain

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2, \text{ and } \|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2$$

By property II specified in Definition 2, we have

$$\|\hat{L}_{k+1}^{11} - L^1\|_2 \leq \|\hat{L}_k^{11} - L^1\|_2, \text{ and } \|\hat{L}_{k+1}^{22} - L^2\|_2 \leq \|\hat{L}_k^{22} - L^2\|_2$$

Then, (3.51) and (3.52) are verified because

$$\|\hat{L}_k^{22} - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2 = \max \left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2 \right)$$

and

$$\|\hat{L}_k^{11} - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2 = \max \left(\|\hat{L}_k^{11} - L^1\|_2, \|\hat{L}_k^{21} - L^1\|_2 \right)$$

Figure 3.8: Three typical cases of the relation between $(\hat{L}_k^{12}, \hat{L}_k^{11})$ and $(\hat{L}_k^{22}, \hat{L}_k^{21})$ for a scalar example.

If Case 2 is true, then by property II in Definition 2,

$$\|\hat{L}_k^{12} - L^2\|_2 < \|\hat{L}_k^{22} - L^2\|_2 \Rightarrow \|f^1(\hat{L}_k^{12}) - L^1\|_2 < \|f^1(\hat{L}_k^{22}) - L^1\|_2$$

where $f^1(\hat{L}_k^{12}) = \hat{L}_k^{11}$ and $f^1(\hat{L}_k^{22}) = f^1(f^2(\hat{L}_k^{21}))$. Using property I specified in Definition 2, we have

$$\|f^1(\hat{L}_k^{22}) - L^1\|_2 = \|f^1(f^2(\hat{L}_k^{21})) - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2$$

So, $\|\hat{L}_k^{11} - L^1\|_2 < \|\hat{L}_k^{21} - L^1\|_2$. From (3.49) and (3.50), we have

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \|\hat{L}_k^{22} - L^2\|_2 = \max\left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2\right) \quad (3.53)$$

$$\|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2 \quad (3.54)$$

Then by property II in Definition 2,

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \|\hat{L}_k^{22} - L^2\|_2 \Rightarrow \|\hat{L}_{k+1}^{11} - L^1\|_2 \leq \|f^1(\hat{L}_k^{22}) - L^1\|_2 \quad (3.55)$$

$$\|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2 \Rightarrow \|\hat{L}_{k+1}^{22} - L^2\|_2 \leq \|\hat{L}_k^{22} - L^2\|_2 \quad (3.56)$$

So, (3.51) is proved from (3.53) and (3.56). In addition, by property I in Definition 2,

$$\|f^1(\hat{L}_k^{22}) - L^1\|_2 = \|f^1(f^2(\hat{L}_k^{21})) - L^1\|_2 \leq \|\hat{L}_k^{21} - L^1\|_2 \quad (3.57)$$

From (3.54), (3.55) and (3.57), we proved (3.52).

If Case 3 is true, we can prove (3.51)-(3.52) via a similar way. By property II in Definition 2,

$$\|\hat{L}_k^{21} - L^1\|_2 < \|\hat{L}_k^{11} - L^1\|_2 \Rightarrow \|f^2(\hat{L}_k^{21}) - L^2\|_2 < \|f^2(\hat{L}_k^{11}) - L^2\|_2$$

where $f^2(\hat{L}_k^{21}) = \hat{L}_k^{22}$ and $f^2(\hat{L}_k^{11}) = f^2(f^1(\hat{L}_k^{12}))$. Using property I specified in Definition 2, we have

$$\|f^2(\hat{L}_k^{11}) - L^2\|_2 = \|f^2(f^1(\hat{L}_k^{12})) - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2$$

So, $\|\hat{L}_k^{22} - L^2\|_2 < \|\hat{L}_k^{12} - L^2\|_2$. From (3.49) and (3.50), we have

$$\|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \|\hat{L}_k^{11} - L^1\|_2 = \max\left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2\right) \quad (3.58)$$

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2 \quad (3.59)$$

Then by property II in Definition 2,

$$\|\hat{L}_{k+1}^{21} - L^1\|_2 \leq \|\hat{L}_k^{11} - L^1\|_2 \Rightarrow \|\hat{L}_{k+1}^{22} - L^2\|_2 \leq \|f^2(\hat{L}_k^{11}) - L^2\|_2 \quad (3.60)$$

$$\|\hat{L}_{k+1}^{12} - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2 \Rightarrow \|\hat{L}_{k+1}^{11} - L^1\|_2 \leq \|\hat{L}_k^{11} - L^1\|_2 \quad (3.61)$$

So, (3.52) is proved from (3.58) and (3.61). In addition, by property I in Definition 2,

$$\|f^2(\hat{L}_k^{11}) - L^2\|_2 = \|f^2(f^1(\hat{L}_k^{12})) - L^2\|_2 \leq \|\hat{L}_k^{12} - L^2\|_2 \quad (3.62)$$

From (3.59), (3.60) and (3.62), we proved (3.51).

After exploring all the three cases, we can conclude that (3.51) and (3.52) are true. The equalities in (3.51) and (3.52) hold only when $e_k^i = 0$. If $\exists K > 0$ such that for $\forall k > K$, $e_k^i \equiv 0$ (WLOG, let $e_k^1 \equiv 0$). Then, provided that the excitation conditions hold, $\hat{L}_k^{12} = \hat{L}_k^{22}$. By Lemma 3, we have $\hat{L}_k^{21} = \hat{L}_k^{11}$. Then

$$\begin{aligned} \|\hat{L}_k^{21} - L^1\|_2 &= \|\hat{L}_k^{11} - L^1\|_2 = \|f^1(\hat{L}_k^{12}) - L^1\|_2 \\ &= \|f^1(\hat{L}_k^{22}) - L^1\|_2 = \|f^1(f^2(\hat{L}_k^{21})) - L^1\|_2 \end{aligned}$$

From property I in Definition 2, we can conclude that $\hat{L}_k^{21} \equiv L^1$. Similarly we can prove $\hat{L}_k^{12} \equiv L^2$.

If e_k^i is not always zero, then from (3.51)-(3.52), we know that

$$\max\left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2\right), \text{ and } \max\left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2\right)$$

are decreasing. Since both are bounded from below by 0,

$$\max \left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2 \right) \longrightarrow 0, \text{ and } \max \left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2 \right) \longrightarrow 0$$

as $k \rightarrow +\infty$. This completes the proof. ■

Remark 12 *Another method to satisfy the excitation condition is to add small white noise to the system. The system dynamics is specified in (3.20) with cost functions (3.21)-(3.22). It is well known that the Nash strategies are the same as in (3.4)-(3.5). We choose the following approximate estimator for player 1 and player 2, respectively:*

$$\hat{L}_k^{12} = \hat{L}_{k-1}^{12} - e_k^1 \phi_k^\top P_k^1 \quad (3.63)$$

$$\hat{L}_k^{21} = \hat{L}_{k-1}^{21} - e_k^2 \phi_k^\top P_k^2 \quad (3.64)$$

where

$$e_k^1 = [B^{2\top} B^2]^{-1} B^{2\top} (\hat{x}_k^1 - x_k) = [\hat{L}_{k-1}^{12} - \hat{L}_{k-1}^{22}] \phi_k - [B^{2\top} B^2]^{-1} B^{2\top} w_{k-1}$$

$$e_k^2 = [B^{1\top} B^1]^{-1} B^{1\top} (\hat{x}_k^2 - x_k) = [\hat{L}_{k-1}^{21} - \hat{L}_{k-1}^{11}] \phi_k - [B^{1\top} B^1]^{-1} B^{1\top} w_{k-1}$$

$$P_k^1 = P_{k-1}^1 - P_{k-1}^1 \phi_k [\lambda^1 + \phi_k^\top P_{k-1}^1 \phi_k]^{-1} \phi_k^\top P_{k-1}^1$$

$$P_k^2 = P_{k-1}^2 - P_{k-1}^2 \phi_k [\lambda^2 + \phi_k^\top P_{k-1}^2 \phi_k]^{-1} \phi_k^\top P_{k-1}^2$$

with the sufficiently large initial conditions $P_0^1 > 0$ and $P_0^2 > 0$. Using the similar method in the proof of Theorem 6, we can prove the following two inequalities: ,

$$\begin{aligned} \max \left(\|E_{w_k}(\hat{L}_{k+1}^{12}) - L^2\|_2, \|E_{w_k}(\hat{L}_{k+1}^{22}) - L^2\|_2 \right) &\leq \max \left(\|\hat{L}_k^{12} - L^2\|_2, \|\hat{L}_k^{22} - L^2\|_2 \right) \\ \max \left(\|E_{w_k}(\hat{L}_{k+1}^{21}) - L^1\|_2, \|E_{w_k}(\hat{L}_{k+1}^{11}) - L^1\|_2 \right) &\leq \max \left(\|\hat{L}_k^{21} - L^1\|_2, \|\hat{L}_k^{11} - L^1\|_2 \right) \end{aligned}$$

where $E_{w_k}(\bullet)$ denotes the mathematical expectation with respect to w_k .

3.3.3 Extension to LQSGs

We can extend the adaptive laws in (3.41)-(3.42) to 2-player infinite-horizon LQSGs defined in (3.24)-(3.27).

In the two-side adaptation scheme specified in the beginning part of Section 3.3.1, when k is odd, player 2 will update $\hat{L}_k^{21}$ according to the adaptive law (3.42), and player 1 will apply the best response controller calculated from the design equation of player 1. When k is even, player 1 will update $\hat{L}_k^{12}$ according to the adaptive law (3.41), and player 2 will apply the best response controller based on its design equation.

Theorem 7 *Consider the two-side adaptation scheme of LQSGs defined in (3.24)-(3.25) with the cost functions (3.18)-(3.19) for the reference input tracking case, the convergent adaptive law for player 1 and player 2 are the same as in (3.41) and (3.42), respectively, except that*

$$\phi_k^1 = \begin{cases} 0 & , \text{ when } k \text{ is even} \\ x_{k-1} & , \text{ when } k \text{ is odd} \end{cases} \quad (3.65)$$

$$\phi_k^2 = \begin{cases} 0 & , \text{ when } k \text{ is odd} \\ x_{k-1} & , \text{ when } k \text{ is even} \end{cases} \quad (3.66)$$

and after getting $\hat{L}_k^{12}$ and $\hat{L}_k^{21}$, we can calculate the best response control gains based on the following design functions:

$$\hat{L}_k^1 = \begin{cases} 0 & , \mathcal{K}_e \\ \hat{L}_k^{11} & , \mathcal{K}_o \end{cases} \quad (3.67)$$

$$\hat{L}_k^2 = \begin{cases} \hat{L}_k^{22} & , \mathcal{K}_e \\ 0 & , \mathcal{K}_o \end{cases} \quad (3.68)$$

with

$$\hat{L}_k^{11} = -[R^{11} + B^{1\top} \hat{K}_{e,k}^1 B^1]^{-1} B^{1\top} \hat{K}_{e,k}^1 A \quad (3.69)$$

$$\hat{L}_k^{22} = -[R^{22} + B^{2\top} \hat{K}_{o,k}^2 B^2]^{-1} B^{2\top} \hat{K}_{o,k}^2 A \quad (3.70)$$

$\hat{K}_{e,k}^1$ and $\hat{K}_{o,k}^2$ are the positive semi-definite solutions to the following two DAREs, respectively,

$$\bar{A}^{1\top} \hat{K}_{e,k}^1 \bar{A}^1 - \hat{K}_{e,k}^1 - \bar{A}^{1\top} \hat{K}_{e,k}^1 \bar{B}^1 [\bar{B}^{1\top} \hat{K}_{e,k}^1 \bar{B}^1 + \bar{R}^1]^{-1} \bar{B}^{1\top} X \bar{A}^1 + \bar{Q}^1 = 0 \quad (3.71)$$

$$\bar{A}^{2\top} \hat{K}_{o,k}^2 \bar{A}^2 - \hat{K}_{o,k}^2 - \bar{A}^{2\top} \hat{K}_{o,k}^2 \bar{B}^2 [\bar{B}^{2\top} \hat{K}_{o,k}^2 \bar{B}^2 + \bar{R}^2]^{-1} \bar{B}^{2\top} X \bar{A}^2 + \bar{Q}^2 = 0 \quad (3.72)$$

where $\bar{Q}^1 = [A + B^2 \hat{L}_k^{12}]^\top Q^1 [A + B^2 \hat{L}_k^{12}] + Q^1 + \hat{L}_k^{12\top} R^{12} \hat{L}_k^{12}$, $\bar{A}^1 = A[A + B^2 \hat{L}_k^{12}]$, $\bar{B}^1 = B^1$, $\bar{R}^1 = R^{11}$, $\bar{Q}^2 = [A + B^1 \hat{L}_k^{21}]^\top Q^2 [A + B^1 \hat{L}_k^{21}] + Q^2 + \hat{L}_k^{21\top} R^{21} \hat{L}_k^{21}$, $\bar{A}^2 = A[A + B^1 \hat{L}_k^{21}]$, $\bar{B}^2 = B^2$, and $\bar{R}^2 = R^{22}$.

Proof: The proof follows the one in Section 3.3.2 and the discussion in Section 3.3.3 (just before this Theorem) and part (iii) in Theorem 5. ■

3.3.4 Simulations and Discussions

Numerical examples are given in this section to illustrate our two-side adaptation for the state feedback Nash strategies of infinite-horizon linear quadratic games.

Linear Quadratic Simultaneous Game Case

Consider the system defined in (3.1) - (3.3) with

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, Q^1 = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}, Q^2 = \begin{bmatrix} 1.6 & 0 \\ 0 & 2.1 \end{bmatrix}$$

$B^1 = [1.3, 1.7]^\top$, $B^2 = [1.5, 1.6]^\top$, $R^{11} = 0.4$, $R^{12} = 1.0$, $R^{21} = 1.0$, and $R^{22} = 0.5$.

We can calculate K^1 , K^2 , L^1 and L^2 from the coupled discrete-time algebraic Riccati equation specified in (3.6)-(3.9),

$$K^1 = \begin{bmatrix} 2.3609 & 1.0957 \\ 1.0957 & 3.4812 \end{bmatrix}, K^2 = \begin{bmatrix} 2.5639 & 1.2059 \\ 1.2059 & 4.2965 \end{bmatrix}$$

$L^1 = [-0.4992, -1.1682]$ and $L^2 = [-0.5215, -1.1295]$. We first carry out the two-side adaptation based on the adaptive laws (3.41)-(3.42) for the system with cost functions in (3.18)-(3.19). The parameters of adaptive laws are chosen as $\Gamma_1 = 0.8I$, $\Gamma_2 = 0.9I$, $\kappa_1 = 1.1$ and $\kappa_2 = 1.7$. The reference signals are generated by

$$r_k = \begin{bmatrix} 1.7 \sin(3.5k) \\ 3.5 \cos(0.9k^2 + 0.4k + 0.1) \end{bmatrix}$$

As shown in Fig. 3.9, we can see that when $k \rightarrow +\infty$,

$$\hat{L}_k^{12} \rightarrow L^2, \quad \hat{L}_k^{22} \rightarrow L^2, \quad \hat{L}_k^{21} \rightarrow L^1, \quad \hat{L}_k^{11} \rightarrow L^1, \quad \hat{x}_k^1 \rightarrow x_k, \quad \text{and} \quad \hat{x}_k^2 \rightarrow x_k.$$

We can verify the relations specified in (3.51)-(3.52) from Fig. 3.10, which shows that $\max(\|\hat{L}_k^{11} - L^1\|_2, \|\hat{L}_k^{21} - L^1\|_2)$ and $\max(\|\hat{L}_k^{22} - L^2\|_2, \|\hat{L}_k^{12} - L^2\|_2)$ are non-increasing sequences.

We also implement the approximate adaptive laws specified in Remark 12 for the system with the white noise

$$w \sim \mathcal{N}(\mu = [0, 0]^\top, \sigma^2 = 0.01I_2)$$

as the system disturbance. The results are shown in Fig. 3.11.

Linear Quadratic Sequential Game Case

Let us consider the LQSG defined in (3.24)-(3.27) with

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, B^1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, B^2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, Q^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

and $R^{11} = 1$, $R^{12} = 2$, $R^{21} = 2$, $R^{22} = 1$. We can calculate K_o^i , K_e^i , L^1 and L^2 from the coupled discrete-time algebraic Riccati equation specified in Remark 3 of Section 2.2,

$$K_o^1 = \begin{bmatrix} 4.9829 & 5.2619 \\ 5.2619 & 8.1701 \end{bmatrix}, K_e^1 = \begin{bmatrix} 19.6977 & 23.2266 \\ 23.2266 & 29.9641 \end{bmatrix}$$

Figure 3.9: $\hat{L}^{ij}$, L^j , x_k and $\hat{x}_k^i$ in two-side adaptation for a simultaneous game: reference signal tracking case. $\{\hat{L}_k^{ij}\}_s$ is the s^{th} element of $\hat{L}_k^{ij}$. The reference signal is $r_k = [1.7 \sin(3.5k), 3.5 \cos(0.9k^2 + 0.4k + 0.1)]^\top$. We can see the convergence of $\hat{L}_k^{ij}$ to L^j .

Figure 3.10: As proved in Theorem 6, $\max(\|\hat{L}_k^{11} - L^1\|_2, \|\hat{L}_k^{21} - L^1\|_2)$ and $\max(\|\hat{L}_k^{22} - L^2\|_2, \|\hat{L}_k^{12} - L^2\|_2)$ are theoretically non-increasing sequences. In our simulation, due to the calculation accuracy, we found several increasing points: $D_{11}^1 = 0.0861$, $D_{12}^1 = 0.0944$, and $D_{11}^2 = 0.0838$, $D_{12}^2 = 0.0931$.

Figure 3.11: $\hat{L}^{ij}$, L^j , x_k and $\hat{x}_k^i$ in the two-side adaptation for a simultaneous game with small system noise. $\{\hat{L}_k^{ij}\}_s$ is the s^{th} element of $\hat{L}_k^{ij}$. The system noise is $w = \mathcal{N}([0, 0]^\top, 0.01I)$. The initial values of parameters are $P_0^1 = P_0^2 = 100I$ and $\lambda^1 = \lambda^2 = 100.0$.

$$K_o^2 = \begin{bmatrix} 9.9557 & 10.5162 \\ 10.5162 & 17.4163 \end{bmatrix}, K_e^2 = \begin{bmatrix} 29.3923 & 32.6710 \\ 32.6710 & 42.0584 \end{bmatrix}$$

$$L^1 = [-1.9914, -2.6310] \text{ and } L^2 = [-2.1400, -2.8267].$$

We first implement the adaptive scheme based on the adaptive law (3.28)-(3.29) with $\Gamma_1 = \Gamma_2 = I$, $\kappa_1 = \kappa_2 = 0.4$ and

$$r_k = \begin{bmatrix} 1.7 \sin(0.5k) \\ \cos(0.9k) \end{bmatrix}$$

As shown in Fig. 3.12, we can see that $\hat{L}_k^{i1}$, $\hat{L}_k^{i2}$, and $\hat{x}_k^i$ converge to L^1 , L^2 and x_k , respectively. We can verify the relations specified in (3.51)-(3.52) from Fig. 3.13, which shows that $\max(\|\hat{L}_k^{11} - L^1\|_2, \|\hat{L}_k^{21} - L^1\|_2)$ and $\max(\|\hat{L}_k^{22} - L^2\|_2, \|\hat{L}_k^{12} - L^2\|_2)$ are non-increasing sequences.

For the same system with the white noise

$$w \sim \mathcal{N}(\mu = [0, 0]^\top, \sigma^2 = 0.01I_2)$$

as the system disturbance, we also implement the approximate adaptive laws specified in Remark 12 with the special ϕ_k^1 and ϕ_k^2 defined in Theorem 7. The results are shown in Fig. 3.14.

Figure 3.12: $\hat{L}^{ij}$, L^j , x_k and $\hat{x}_k^i$ in the two-side adaptation for a LQSG: reference signal tracking case. $\{\hat{L}_k^{ij}\}_s$ is the s^{th} element of $\hat{L}_k^{ij}$. The reference signal is $r_k = [1.7 \sin(0.5k), \cos(0.9k)]^\top$. We can see the convergence of $\hat{L}_k^{ij}$ to L^j .

Figure 3.13: As proved in Theorem 6, $\max(\|\hat{L}_k^{11} - L^1\|_2, \|\hat{L}_k^{21} - L^1\|_2)$ and $\max(\|\hat{L}_k^{22} - L^2\|_2, \|\hat{L}_k^{12} - L^2\|_2)$ are theoretically non-increasing sequences. In our simulation, due to the calculation accuracy when small values are involved, we found one increasing points: $D_{14}^2 = 0.0179$, $D_{15}^1 = 0.0198$. 86

Figure 3.14: $\hat{L}^{ij}$, L^j , x_k and $\hat{x}_k^i$ in the two-side adaptation for a LQSG with small system noise. $\{\hat{L}_k^{ij}\}_s$ is the s^{th} element of $\hat{L}_k^{ij}$. The system noise is $w = \mathcal{N}([0, 0]^\top, 0.01I)$. The initial values of parameters are $P_0^1 = P_0^2 = 1000I$ and $\lambda^1 = \lambda^2 = 2.0$.

3.4 Chapter Summary

In this chapter, we focused on the adaptivity analysis of the infinite-horizon perfect-state two-player linear quadratic games, including simultaneous and sequential ones, under a FP framework. We first considered the adaptive control issues for simultaneous linear quadratic game in a one-side adaptation scheme, in which only one of the two players has perfect information of the cost functions of both players, i.e., the player knows exact Q^1 , Q^2 , R^{11} , R^{12} , R^{21} , R^{22} , while the other player only has access to its cost function and does not know the parameters of the cost function of its opponent. We assumed that both players will apply Nash strategies based on the available information set. We also assumed that the player without perfect information knows in advance that its opponent will implement Nash strategies, based on which it can design his adaptive control. We designed an indirect adaptive method to estimate the control gain of the player with perfect information. With stabilizability and detectability conditions, we proved that given the estimated control gain, the player with imperfect information can always calculate his associated control gain from the coupled discrete-time algebraic Riccati equations. Convergence was proved and the excitation conditions were satisfied as well. We extended our approach to LQSGs via specially designed ϕ_k .

In the second part of this chapter, adaptivity analysis was considered in a more useful and practical two-side adaptation case, in which each player needs to estimate the control gain of its opponent. We assumed that each player has access to its own cost function as well as the system dynamics, and does not know the parameters of the cost function of its opponent. i.e., player 1 knows Q^1 , R^{12} , R^{11} , A , B^1 and B^2 while player 2 has access to Q^2 , R^{21} , R^{22} , A , B^1 and B^2 . We also assumed that each player

knows in advance that its opponent will implement Nash strategies. So, player 2 (or player 1) will estimate the control gain L^1 (or L^2) first, and then calculate/estimate its own control gain L^2 (or L^1). We provided adaptation design, parameter convergence proof, and persistence of excitation for the two-side adaptation scheme. Similar to the one-side adaptation scheme, we extended the adaptive law and convergence proof to LQSGs.

Finally, numerical examples were simulated to illustrate our adaptive approaches for one-side and two-side adaptive schemes.

CHAPTER 4

DISTURBANCE ATTENUATION ANALYSIS OF STATE FEEDBACK NASH STRATEGY FOR LQSG

In this chapter, we use a zero-sum game theoretical approach to address the disturbance attenuation analysis of our state feedback Nash strategies for LQSGs with uncertainties. A H^∞ –optimal controller for each player is designed in the context of sensitivity reduction.

4.1 Introduction

During the past decades, tremendous effort has been devoted to robust control problems. Between 1942 and 1963, Black [60], Nyquist [61], Bode [62] and Horowitz [63] used classical sensitivity methods to solve the Single Input Single Output (SISO) robust control problem by focusing on sensitivity reduction and noise suppression. However, it was not until the result of Cruz and Perkins [64] in 1964 that the Multiple Input Multiple Output (MIMO) uncertainty problem was analyzed by using the “sensitivity comparison matrix”. The current attention on robustness is the preservation of stability only. In 1980, Cruz *et al* [65] demonstrated a relationship between sensitivity and robustness. In the early 1980’s, Zames and Francis [66] used a H-infinity norm to reduce sensitivity in linear plants. Some relevant important

papers were those of Zames [67], Zames and Bensoussan [68], Doyle *et al* [69] and Stoorvogel *et al* [70].

Due to the uncertainties in the robust control problem, a game theoretic or minimax approach was advocated by Ragade and Sarma [71], Dorato and Kestenbaum [72], Salmon [73], Başar and Bernhard [38], and Witsenhausen [74]. In this approach, the uncertain parameters are controlled by a virtual player which maximizes the cost function that is being minimized by the original control variables.

The main contribution of this chapter is to address the disturbance attenuation analysis of our state feedback Nash strategies for LQSGs with uncertainties or disturbances, which are modeled by a virtual player in a zero-sum game. For finite-horizon LQSGs, we provide state feedback Nash strategies with the optimal attenuation level. We extend the approach to infinite-horizon LQSGs. We prove that the feedback system is Bounded Input Bounded Output (BIBO) stable with respect to the disturbances.

The remainder of this chapter is organized as follows. Section 4.2 reviews the mathematical relation between zero-sum games and H^∞ -optimal control problems proposed by Başar and Bernhard [38]. In Section 4.3, we first explore the relation between zero-sum games and LQSGs with system disturbances. Secondly, we solve the related zero-sum games in a finite-horizon scheme. State feedback Nash strategies are provided. Existence condition are presented in the sense that for any disturbance the upper values of the related zero-sum games are bounded. Thirdly, a optimal disturbance attenuation level $\sqrt{\hat{\lambda}}$ is defined for the original LQSGs. We also extend our approach to infinite-horizon LQSGs. We prove that the closed-loop system with the state feedback Nash strategies is Bounded Input Bounded Output (BIBO) stable

with respect to the disturbances. Numerical examples are presented in Section 4.4 to illustrate the disturbance attenuation oriented state feedback Nash strategies for LQSGs with system disturbances. Finally, we draw conclusions in Section 4.5.

4.2 H^∞ -optimal Control and Related Zero-sum Games

Consider the worst-case controller design problem (see [38]):

$$\left. \begin{aligned} z &= G_{11}w + G_{12}u \\ y &= G_{21}w + G_{22}u \\ u &= \gamma(y) \end{aligned} \right\} \Sigma$$

In this set-up, w, u, z and y denote, respectively, the exogenous disturbance, the control variables, the controlled output such as tracking errors and filtered actuator signals, and the measured output. They belong to the appropriate Hilbert spaces, denoted as $\mathcal{H}_w$, $\mathcal{H}_u$, $\mathcal{H}_z$, and $\mathcal{H}_y$, respectively. The transfer matrix G and control strategy $\gamma(\bullet)$ are, by assumption, real-rational and proper.

Given some regularity conditions, for any fixed control strategy $u = \gamma(y)$, $\exists$ a mapping $T_\gamma: \mathcal{H}_w \rightarrow \mathcal{H}_z$. The design objective is to find an optimal control strategy $u = \gamma^*(y)$ to minimize the H_∞ -norm of the corresponding mapping T_{γ^*} , i.e.,

$$\gamma^* := \arg_\gamma \left\{ \inf_{\gamma \in \mathcal{M}} \sup_{w \in \mathcal{H}_w} \|T_\gamma(w)\|_z / \|w\|_w \right\} \quad (4.1)$$

Assume that $\exists \gamma^*$ which satisfies the above equation (4.1) with the corresponding optimal value $\sqrt{\lambda^*}$, which, in mathematical terms, is

$$\sqrt{\lambda^*} = \sup_{w \in \mathcal{H}_w} \|T_{\gamma^*}(w)\|_z / \|w\|_w \leq \sup_{w \in \mathcal{H}_w} \|T_\gamma(w)\|_z / \|w\|_w \quad (4.2)$$

We can equivalently rewrite equation (4.2) into:

$$1) \forall w \in \mathcal{H}_w, \|T_{\gamma^*}(w)\|_z^2 \leq \lambda^* \|w\|_w^2$$

2) there exists no $\hat{\gamma} \neq \gamma^*$ and $\hat{\lambda} < \lambda^*$, s.t., $\forall w \in \mathcal{H}_w$, $\|T_{\hat{\gamma}}(w)\|_z^2 \leq \hat{\lambda}\|w\|_w^2$

By introducing the parameterized (in $\lambda \geq 0$) family of cost functions:

$$J_\lambda(\gamma, w) = \|T_\gamma\|_z^2 - \lambda\|w\|_w^2 \equiv J(\gamma, w) - \lambda\|w\|_w^2 \quad (4.3)$$

the H_∞ -optimal control problem is translated into a related zero-sum game problem, in which the cost function (4.3) is being minimized by the control variables and maximized by the disturbance or uncertainty.

4.3 Disturbance Attenuation Analysis of LQSGs with State Feedback Nash Strategies

4.3.1 Zero-sum Games for Disturbance Attenuation Analysis of LQSGs

We consider the following 2-player LQSGs with disturbances:

$$x_{k+1} = A_k x_k + B_k^1 u_k^1 + B_k^2 u_k^2 + D_k w_k \quad (4.4)$$

with the constraints

$$\begin{cases} u_k^1 = 0, & \text{when } k \text{ is even} \\ u_k^2 = 0, & \text{when } k \text{ is odd} \end{cases} \quad (4.5)$$

where, $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}^{r_i}$, $1 \leq i \leq 2$, $0 \leq k \leq N-1$.

The cost functions of the players are defined by

$$J^1 = x_N^\top Q_N^1 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + u_k^{2\top} R_k^{12} u_k^2) \quad (4.6)$$

$$J^2 = x_N^\top Q_N^2 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^2 x_k + u_k^{1\top} R_k^{21} u_k^1 + u_k^{2\top} R_k^{22} u_k^2) \quad (4.7)$$

where all matrices are symmetric with $Q_k^i \geq 0$ and $R_k^{ij} > 0$ for $i = 1, 2$, $j = 1, 2$, and $k = 0, 1, \dots, N$.

Similarly, the related zero-sum games have the following parameterized cost functions:

$$J_\lambda^1 = x_N^\top Q_N^1 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + u_k^{2\top} R_k^{12} u_k^2 - \delta_k^1 \lambda w_k^\top w_k) \quad (4.8)$$

$$J_\lambda^2 = x_N^\top Q_N^2 x_N + \sum_{k=0}^{N-1} (x_k^\top Q_k^2 x_k + u_k^{1\top} R_k^{21} u_k^1 + u_k^{2\top} R_k^{22} u_k^2 - \delta_k^2 \lambda w_k^\top w_k) \quad (4.9)$$

where $\lambda > 0$ and

$$\delta_k^1 = \begin{cases} 0, & k \text{ is even} \\ 1, & k \text{ is odd} \end{cases}, \delta_k^2 = \begin{cases} 1, & k \text{ is even} \\ 0, & k \text{ is odd} \end{cases} \quad (4.10)$$

At each odd-valued time k , player 1 (u_k^1) will minimize its parameterized cost function J_λ^1 and disturbance (w_k) will try to maximize J_λ^1 . On the other hand, when k is even, J_λ^2 will be minimized by the control variable u_k^2 and maximized by the disturbance or uncertainty w_k .

Remark 13 *Here we assume that at each time k , the disturbance w_k is trying to maximize the cost function of the player who applies its controller.*

4.3.2 State Feedback Nash Strategies for Finite-Horizon LQSGs with Soft-Constrained Cost Functions

In this section, we will derive the state feedback Nash Strategies for the LQSGs with system dynamics (4.4)-(4.5) and soft-constrained cost functions (4.8)-(4.9).

Theorem 8 *For a LQSG defined in (4.4)-(4.5) and (4.8)-(4.9) with a fixed $\lambda > 0$,*

(i) *There exists a unique state feedback Nash strategies if, and only if, $\Xi_k > 0$.*

Ξ_k is defined as

$$\Xi_k = \begin{cases} \lambda I - D_k^\top K_{k+1}^2 D_k, & k \text{ is even} \\ \lambda I - D_k^\top K_{k+1}^1 D_k, & k \text{ is odd} \end{cases} \quad (4.11)$$

(Ξ_k is not necessarily symmetric. $\Xi_k > 0$ means that the symmetric part $\frac{1}{2}[\Xi_k + \Xi_k^\top]$ is positive definite. See [75]) where

$$K_k^1 = \begin{cases} Q_k^1 + A_k^\top [\Lambda_k^{-1}]^\top [K_{k+1}^1 + \\ K_{k+1}^{2\top} B_k^2 (R_k^{22})^{-1} R_k^{12} (R_k^{22})^{-1} B_k^{2\top} K_{k+1}^2] \Lambda_k^{-1} A_k & , \quad k \text{ is even} \\ Q_k^1 + A_k^\top K_{k+1}^1 \Lambda_k^{-1} A_k & , \quad k \text{ is odd} \end{cases} \quad (4.12)$$

$$K_k^2 = \begin{cases} Q_k^2 + A_k^\top K_{k+1}^2 \Lambda_k^{-1} A_k & , \quad k \text{ is even} \\ Q_k^2 + A_k^\top [\Lambda_k^{-1}]^\top [K_{k+1}^2 + \\ K_{k+1}^{1\top} B_k^1 (R_k^{11})^{-1} R_k^{21} (R_k^{11})^{-1} B_k^{1\top} K_{k+1}^1] \Lambda_k^{-1} A_k & , \quad k \text{ is odd} \end{cases} \quad (4.13)$$

with boundary conditions $K_N^i = Q_N^i$. The Λ_k is defined as

$$\Lambda_k = \begin{cases} I + [B_k^2 (R_k^{22})^{-1} B_k^{2\top} - \frac{1}{\lambda} D_k D_k^\top] K_{k+1}^2, & k \text{ is even} \\ I + [B_k^1 (R_k^{11})^{-1} B_k^{1\top} - \frac{1}{\lambda} D_k D_k^\top] K_{k+1}^1, & k \text{ is odd} \end{cases} \quad (4.14)$$

(ii) Under the condition $\Xi_k > 0$ (Ξ_k defined in (4.11)), then the matrix Λ_k is invertible, and the state feedback strategies are

$$\gamma_k^1 = u_k^{1*} = \begin{cases} 0 & , \quad k \text{ is even} \\ -(R_k^{11})^{-1} B_k^{1\top} K_{k+1}^1 \Lambda_k^{-1} A_k x_k & , \quad k \text{ is odd} \end{cases} \quad (4.15)$$

$$\gamma_k^2 = u_k^{2*} = \begin{cases} -(R_k^{22})^{-1} B_k^{2\top} K_{k+1}^2 \Lambda_k^{-1} A_k x_k & , \quad k \text{ is even} \\ 0 & , \quad k \text{ is odd} \end{cases} \quad (4.16)$$

$$\gamma_k^w = w_k^* = \begin{cases} \frac{1}{\lambda} D_k^\top K_{k+1}^2 \Lambda_k^{-1} A_k x_k & , \quad k \text{ is even} \\ \frac{1}{\lambda} D_k^\top K_{k+1}^1 \Lambda_k^{-1} A_k x_k & , \quad k \text{ is odd} \end{cases} \quad (4.17)$$

and the optimal cost functions are

$$J_\lambda^{1*} = x_0^\top K_0^1 x_0$$

$$J_\lambda^{2*} = x_0^\top K_0^2 x_0$$

where x_0 is the initial state of the system defined in (4.4).

(iii) If the matrix Ξ_k in (4.11) has a negative eigenvalue for some $k \in [0, N-1]$, then the LQSG does not admit a Nash strategy under any information structure, and the upper value of the LQSG becomes unbounded.

Proof: We first prove the sufficient part of (i) and the related part (ii). Using dynamic programming, we define

$$G_k^1 = x_k^\top Q_k^1 x_k + u_k^{1\top} R_k^{11} u_k^1 + \gamma_k^{2\top} R_k^{12} \gamma_k^2 - \delta_k^1 \lambda w_k^\top w_k + V_{k+1}^1 \quad (4.18)$$

$$G_k^2 = x_k^\top Q_k^2 x_k + u_k^{2\top} R_k^{22} u_k^2 + \gamma_k^{1\top} R_k^{21} \gamma_k^1 - \delta_k^2 \lambda w_k^\top w_k + V_{k+1}^2 \quad (4.19)$$

$$V_k^1 = \min_{u_k^1} \max_{w_k} G_k^1 = \max_{w_k} \min_{u_k^1} G_k^1 \quad (4.20)$$

$$V_k^2 = \min_{u_k^2} \max_{w_k} G_k^2 = \max_{w_k} \min_{u_k^2} G_k^2 \quad (4.21)$$

and $V_N^i(x_N) = x_N^\top Q_N^i x_N$. It is clear that the optimal costs $J_\lambda^{1*} = V_0^1$ and $J_\lambda^{2*} = V_0^2$. We will verify the second equalities in (4.20)-(4.21) and the conclusion of $V_k^i = x_k^\top K_k^i x_k$ in the following proof.

When $k = N - 1$ and $N - 1$ is odd, we have

$$G_{N-1}^1 = x_{N-1}^\top Q_{N-1}^1 x_{N-1} + u_{N-1}^{1\top} R_{N-1}^{11} u_{N-1}^1 - \lambda w_{N-1}^\top w_{N-1} + x_N^\top Q_N^1 x_N$$

because $u_{N-1}^2 = 0$, $\delta_{N-1}^1 = 1$, and $V_N^1 = x_N^\top Q_N^1 x_N$. G_{N-1}^1 is strictly convex in u_{N-1}^1 due to the fact that $R_{N-1}^{11} + B_{N-1}^{1\top} Q_N^1 B_{N-1}^1 > 0$. Since $\Xi_{N-1} = \lambda I - D_{N-1}^\top Q_{N-1}^1 D_{N-1} > 0$, G_{N-1}^1 is strictly concave in w_{N-1} . Then, there are unique $\gamma_{N-1}^w = w_{N-1}^*$ and $\gamma_{N-1}^1 = u_{N-1}^{1*}$ to make V_{N-1}^1 optimal. The optimal strategies can be uniquely obtained by setting

$$\begin{aligned} \frac{\partial G_{N-1}^1}{\partial u_{N-1}^1} &= R_{N-1}^{11} u_{N-1}^1 + B_{N-1}^{1\top} Q_N^1 x_N = 0 \\ \frac{\partial G_{N-1}^1}{\partial w_{N-1}} &= -\lambda w_{N-1} + D_{N-1}^\top Q_N^1 x_N = 0 \end{aligned}$$

So

$$\gamma_{N-1}^1 = -(R_{N-1}^{11})^{-1} B_{N-1}^{1\top} Q_N^1 x_N, \quad \gamma_{N-1}^w = \frac{1}{\lambda} D_{N-1}^\top Q_N^1 x_N$$

Applying them to the system dynamics in (4.4), we obtain

$$\begin{aligned}
x_N &= A_{N-1}x_{N-1} - B_{N-1}^1(R_{N-1}^{11})^{-1}B_{N-1}^{1\top}Q_N^1x_N + \frac{1}{\lambda}D_{N-1}D_{N-1}^\top Q_N^1x_N \\
&= A_{N-1}x_{N-1} - [B_{N-1}^1(R_{N-1}^{11})^{-1}B_{N-1}^{1\top} - \frac{1}{\lambda}D_{N-1}D_{N-1}^\top]Q_N^1x_N \\
&\Rightarrow \\
\Lambda_{N-1}x_N &= A_{N-1}x_{N-1}
\end{aligned}$$

Since the pair of strategies $(\gamma_{N-1}^1, \gamma_{N-1}^2)$ is the unique solution to (4.18)-(4.21) with constrains (4.5) at time $N-1$, there necessarily exist a unique relationship between x_{N-1} and x_N , implying that the matrix Λ_{N-1} should be invertible. So, $x_N = \Lambda_{N-1}^{-1}A_{N-1}x_{N-1}$. Then we can rewrite the optimal strategies as

$$\gamma_{N-1}^1 = -(R_{N-1}^{11})^{-1}B_{N-1}^{1\top}Q_N^1\Lambda_{N-1}^{-1}A_{N-1}x_{N-1} \quad (4.22)$$

$$\gamma_{N-1}^w = \frac{1}{\lambda}D_{N-1}^\top Q_N^1\Lambda_{N-1}^{-1}A_{N-1}x_{N-1} \quad (4.23)$$

Applying the optimal controllers to G_{N-1}^1 , we have

$$\begin{aligned}
V_{N-1}^1 &= x_{N-1}^\top Q_{N-1}^1 x_{N-1} \\
&+ x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^1 B_{N-1}^1 (R_{N-1}^{11})^{-1} B_{N-1}^{1\top} Q_N^1 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \\
&- \frac{1}{\lambda} x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^1 D_{N-1} D_{N-1}^\top Q_N^1 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \\
&+ x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^1 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \\
&= x_{N-1}^\top Q_{N-1}^1 x_{N-1} + x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^1 [I + \\
&\quad (B_{N-1}^1 (R_{N-1}^{11})^{-1} B_{N-1}^{1\top} - \frac{1}{\lambda} D_{N-1} D_{N-1}^\top) Q_N^1] \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}
\end{aligned}$$

Since $\Lambda_{N-1} = I + (B_{N-1}^1 (R_{N-1}^{11})^{-1} B_{N-1}^{1\top} - \frac{1}{\lambda} D_{N-1} D_{N-1}^\top) Q_N^1$ and $V_{N-1}^1 = V_{N-1}^{1\top}$ (for V_{N-1}^1 is a scalar), we can simplify V_{N-1}^1 as

$$\begin{aligned}
V_{N-1}^1 &= x_{N-1}^\top [Q_{N-1}^1 + A_{N-1}^\top Q_N^1 \Lambda_{N-1}^{-1} A_{N-1}] x_{N-1} \\
&= x_{N-1}^\top K_{N-1}^1 x_{N-1}
\end{aligned}$$

On the other hand,

$$G_{N-1}^2 = x_{N-1}^\top Q_{N-1}^2 x_{N-1} + \gamma_{N-1}^{1\top} R_{N-1}^{21} \gamma_{N-1}^1 + x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^2 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}$$

because $u_{N-1}^2 = 0$, $\delta_{N-1}^2 = 0$, and $x_{N-1} = \Lambda_{N-1}^{-1} A_{N-1} x_{N-1}$. Since $u_{N-1}^2 = 0$, $V_{N-1}^2 = G_{N-1}^2$ (From Remark 13, we know that at even time, $V_k^1 = G_k^1$, and $V_k^2 = G_k^2$ for odd time step),

$$\begin{aligned} V_{N-1}^2 &= x_{N-1}^\top Q_{N-1}^2 x_{N-1} \\ &+ x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^1 B_{N-1}^1 (R_{N-1}^{11})^{-1} R_{N-1}^{21} (R_{N-1}^{11})^{-1} B_{N-1}^{1\top} Q_N^1 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \\ &+ x_{N-1}^\top A_{N-1}^\top [\Lambda_{N-1}^{-1}]^\top Q_N^2 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \\ &= x_{N-1}^\top K_{N-1}^2 x_{N-1} \end{aligned}$$

Similarly, when $N - 1$ is even, we can obtain

$$\gamma_{N-1}^2 = -(R_{N-1}^{22})^{-1} B_{N-1}^{2\top} Q_N^2 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \quad (4.24)$$

$$\gamma_{N-1}^w = \frac{1}{\lambda} D_{N-1}^\top Q_N^2 \Lambda_{N-1}^{-1} A_{N-1} x_{N-1} \quad (4.25)$$

and $V_{N-1}^i = x_{N-1}^\top K_{N-1}^i x_{N-1}$. By induction, we assume that $V_k^i = x_k^\top K_k^i x_k$ for $k = l+1$ (already verified for $l = N - 2$). WLOG, let l be even-valued, then,

$$G_l^2 = x_l^\top Q_l^2 x_l + u_l^{2\top} R_l^{22} u_l^2 - \lambda w_l^\top w_l + x_{l+1}^\top K_{l+1}^2 x_{l+1}$$

because $u_l^1 = 0$, $\delta_l^2 = 1$, and $V_{l+1}^2 = x_{l+1}^\top K_{l+1}^2 x_{l+1}$. Since $\Xi_l = \lambda I - D_l^\top K_l^2 D_l > 0$, G_l^2 is strictly concave in w_l . Then, there are unique $\gamma_l^w = w_l^*$ and $\gamma_l^2 = u_l^{2*}$ to make V_l^2 optimal. The optimal strategies can be uniquely obtained by setting

$$\begin{aligned} \frac{\partial G_l^2}{\partial u_l^2} &= R_l^{22} u_l^2 + B_l^{2\top} K_{l+1}^2 x_{l+1} = 0 \\ \frac{\partial G_l^2}{\partial w_l} &= -\lambda w_l + D_l^\top K_{l+1}^2 x_{l+1} = 0 \end{aligned}$$

So

$$\gamma_l^2 = -(R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2 x_{l+1} \quad , \quad \gamma_l^w = \frac{1}{\lambda} D_l^\top K_{l+1}^2 x_{l+1}$$

Applying them to the system dynamics in (4.4), we obtain

$$\begin{aligned} x_{l+1} &= A_l x_l - B_l^2 (R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2 x_{l+1} + \frac{1}{\lambda} D_l D_l^\top K_{l+1}^2 x_{l+1} \\ &= A_l x_l - [B_l^2 (R_l^{22})^{-1} B_l^{2\top} - \frac{1}{\lambda} D_l D_l^\top] K_{l+1}^2 x_{l+1} \\ &\Rightarrow \\ \Lambda_l x_{l+1} &= A_l x_l \end{aligned}$$

Since the pair of strategies (γ_l^1, γ_l^2) is the unique solution to (4.18)-(4.21) with constraints (4.5) at time l , there necessarily exist a unique relationship between x_l and x_{l+1} , implying that the matrix Λ_l should be invertible. So, $x_{l+1} = \Lambda_l^{-1} A_l x_l$. Then we can rewrite the optimal strategies as

$$\gamma_l^2 = -(R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2 \Lambda_l^{-1} A_l x_l \quad (4.26)$$

$$\gamma_l^w = \frac{1}{\lambda} D_l^\top K_{l+1}^2 \Lambda_l^{-1} A_l x_l \quad (4.27)$$

Applying the optimal controllers to G_l^2 , we have

$$\begin{aligned} V_l^2 &= x_l^\top Q_l^2 x_l + x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{2\top} B_l^2 (R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2 \Lambda_l^{-1} A_l x_l \\ &\quad - \frac{1}{\lambda} x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{2\top} D_l D_l^\top K_{l+1}^2 \Lambda_l^{-1} A_l x_l \\ &\quad + x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{2\top} \Lambda_l^{-1} A_l x_l \\ &= x_l^\top Q_l^2 x_l + \\ &\quad x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{2\top} [I + (B_l^2 (R_l^{22})^{-1} B_l^{2\top} - \frac{1}{\lambda} D_l D_l^\top) K_{l+1}^2] \Lambda_l^{-1} A_l x_l \end{aligned}$$

Since $\Lambda_l = I + (B_l^2(R_l^{22})^{-1}B_l^{2\top} - \frac{1}{\lambda}D_l D_l^\top)K_{l+1}^2$ and $V_l^2 = V_l^{2\top}$ (for V_l^2 is a scalar), we can simplify V_l^2 as

$$\begin{aligned} V_l^2 &= x_l^\top [Q_l^2 + A_l^\top K_{l+1}^2 \Lambda_l^{-1} A_l] x_l \\ &= x_l^\top K_l^2 x_l \end{aligned}$$

On the other hand,

$$G_l^1 = x_l^\top Q_l^1 x_l + \gamma_l^{2\top} R_l^{12} \gamma_l^2 + x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{1\top} \Lambda_l^{-1} A_l x_l$$

because $u_l^1 = 0$, $\delta_l^1 = 0$, and $x_{l+1} = \Lambda_l^{-1} A_l x_l$. Since $u_l^1 = 0$, $V_l^1 = G_l^1$.

$$\begin{aligned} V_l^1 &= x_l^\top Q_l^1 x_l \\ &\quad + x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^{2\top} B_l^2 (R_l^{22})^{-1} R_l^{12} (R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2 \Lambda_l^{-1} A_l x_l \\ &\quad + x_l^\top A_l^\top [\Lambda_l^{-1}]^\top K_{l+1}^1 \Lambda_l^{-1} A_l x_l \\ &= x_l^\top \{Q_l^1 + A_l^\top [\Lambda_l^{-1}]^\top [K_{l+1}^1 + K_{l+1}^{2\top} B_l^2 (R_l^{22})^{-1} R_l^{12} (R_l^{22})^{-1} B_l^{2\top} K_{l+1}^2] \Lambda_l^{-1} A_l\} x_l \\ &= x_l^\top K_l^1 x_l \end{aligned}$$

we complete the proof of sufficient part of (i) and part (ii).

Next we will prove part (iii), which is also the necessary part of (i). Contrary to the statement of the part (iii), we suppose that there exists strategies for player 1 and player 2, denoted by μ_k^1 and μ_k^2 , under which, the cost functions J_λ^1 and J_λ^2 are bounded for all possible w_k , even though Ξ_k has at least one negative eigenvalue for some $k \in [0, N-1]$. Let $\bar{k}$ be the largest integer in $[0, N-1]$ for which $\Xi_{\bar{k}}$ has at least one negative eigenvalue, and $\Xi_{\hat{k}} > 0$ for $\forall \hat{k} > \bar{k}$. Furthermore, let a strategy for w_k be chosen as

$$\mu_k^w = \begin{cases} 0 & , \text{ for } k < \bar{k} \\ \bar{w}_{\bar{k}} & , \text{ for } k = \bar{k} \\ \gamma_k^w & , \text{ for } k > \bar{k} \end{cases}$$

where $\bar{w}_{\bar{k}}$ is a design strategy that can be specified arbitrarily, and γ_k^w is defined in (4.17). Let $\mu_{[k_1, k_2]}^1$, $\mu_{[k_1, k_2]}^2$ and $w_{[k_1, k_2]}$ denote the action set which includes actions of the player 1, player 2 and the disturbance, respectively, between time k_1 and k_2 . Denote the state trajectory associated with the action pair $(\mu_{[0, N-1]}^1, \mu_{[0, N-1]}^2, \mu_{[0, N-1]}^w)$ by $\hat{x}_{[0, N]}$. Finally, we introduced the notation $J_\lambda^i(u_{[k, N-1]}^1, u_{[k, N-1]}^2, w_{[k, N-1]}; x_k)$, $i = 1, 2$ to denote the cost functions of a LQSG, formulated in exactly the same way as the original game, with the only difference being that it is defined on time $[k, N-1]$ and has the initial state $\hat{x}_k$. Let

$$\hat{\mu}_k^i = \begin{cases} \mu_k^i & , \text{ for } k \leq \bar{k} \\ \gamma_k^i & , \text{ for } k > \bar{k} \end{cases}$$

Under strategies $\hat{\mu}_k^i$, the cost functions J_λ^1 and J_λ^2 are also bounded for all possible w_k . WLOG, we suppose $\bar{k}$ is odd. Then we have

$$\begin{aligned} & \sup_{w_{[0, N-1]}} J_\lambda^1(\hat{\mu}_{[0, N-1]}^1, \hat{\mu}_{[0, N-1]}^2, w_{[0, N-1]}) \geq J_\lambda^1(\hat{\mu}_{[0, N-1]}^1, \hat{\mu}_{[0, N-1]}^2, \mu_{[0, N-1]}^w) \\ &= J_\lambda^1(\gamma_{[\bar{k}+1, N-1]}^1, \gamma_{[\bar{k}+1, N-1]}^2, \gamma_{[\bar{k}+1, N-1]}^w) \\ & \quad + \sum_{k=0}^{\bar{k}} \{ \hat{x}_k^\top Q_k^1 \hat{x}_k + \mu_k^{1\top} R_k^{11} \mu_k^1 + \mu_k^{2\top} R_k^{12} \mu_k^2 \} - \lambda \bar{w}_{\bar{k}}^\top \bar{w}_{\bar{k}} \\ & \geq \hat{x}_{\bar{k}+1}^\top K_{\bar{k}+1}^1 \hat{x}_{\bar{k}+1} - \lambda \bar{w}_{\bar{k}}^\top \bar{w}_{\bar{k}} \end{aligned}$$

In the above, the inequality in the first line follows that μ_k^w is not necessarily a maximizing strategy for disturbance. The equality of the next line holds because μ_k^w and γ_k^w are the same, $\hat{\mu}_k^i$ and γ_k^i are the same, for $k \in [\bar{k}+1, N-1]$. The last equality comes from the fact that $R_k^{11} > 0$, $R_k^{12} > 0$, and $Q_k^1 \geq 0$.

Since $\Xi_{\bar{k}} = \lambda I - D_{\bar{k}}^\top K_{\bar{k}+1}^1 D_{\bar{k}}$ has at least one negative eigenvalue, then

$$\hat{x}_{\bar{k}+1}^\top K_{\bar{k}+1}^1 \hat{x}_{\bar{k}+1} - \lambda \bar{w}_{\bar{k}}^\top \bar{w}_{\bar{k}}$$

is not bounded above (in $\bar{w}_{\bar{k}}$). Similarly, if $\bar{k}$ is even,

$$\sup_{w_{[0, N-1]}} J_{\lambda}^2(\hat{\mu}_{[0, N-1]}^1, \hat{\mu}_{[0, N-1]}^2, w_{[0, N-1]}) \geq \hat{x}_{\bar{k}+1}^{\top} K_{\bar{k}+1}^2 \hat{x}_{\bar{k}+1} - \lambda \bar{w}_{\bar{k}}^{\top} \bar{w}_{\bar{k}}$$

is not bounded from above either. So a contradiction to the initial hypothesis is found. This completes out proof of part (iii) and the necessity part of (i). $\blacksquare$

Remark 14 In [38], Başar and Bernhard proved that if Ξ_k has a zero (but no negative) eigenvalue, then whether the corresponding zero-sum game at time k admits a saddle point or not depends on the value of the initial state x_0 . Similarly, if $x_0 = 0$, we can allow Ξ_k to have a zero eigenvalue and still preserve the existence of a saddle point.

4.3.3 Disturbance Attenuation Analysis for Finite-Horizon LQSGs with State Feedback Nash Strategies

In this section, we will solve, through a mechanism similar to that introduced in Section 4.2, the discrete-time disturbance attenuation problems for finite-horizon LQSGs with state feedback Nash strategies.

Note that to solve the disturbance attenuation problem, we need to find a closed-loop perfect state feedback Nash strategy pair $\psi = (u_k^1, u_k^2) \in \mathcal{M}_{CLPS}$ ($\mathcal{M}_{CLPS}$ is the action spaces in closed-loop perfect state information structure) such that

$$\sqrt{\lambda^*} = \inf_{\psi} \ll T_{\psi} \gg = \ll T_{\psi^*} \gg \quad (4.28)$$

where $\sqrt{\lambda^*}$ is the optimal attenuation level, and the operator norm of T_{ψ} is

$$\ll T_{\psi} \gg = \sup_{w \in \mathcal{H}_w} \|T_{\psi}(w)\| / \|w\| \equiv \sup_{w \in \mathcal{H}_w} [J^1(\psi, w) + J^2(\psi, w)]^{\frac{1}{2}} / \|w\| \quad (4.29)$$

where J^1 and J^2 are defined in (4.6)-(4.7). To obtain the optimal Nash strategy ψ^* , and the corresponding attenuation level $\sqrt{\lambda^*}$, we first fix λ such that the condition (4.11) $\Xi_k > 0$ is satisfied. For a fixed $\lambda > 0$, let γ_λ^1 and γ_λ^2 denote the associated state feedback strategies generated by (4.15) - (4.16), respectively, and γ_λ^w denote the disturbance policy calculated in (4.17). From Theorem 8, we know that

$$J_\lambda^1(\gamma_\lambda^1, \gamma_\lambda^2, w) + J_\lambda^2(\gamma_\lambda^1, \gamma_\lambda^2, w) \leq J_\lambda^1(\gamma_\lambda^1, \gamma_\lambda^2, \gamma_\lambda^w) + J_\lambda^2(\gamma_\lambda^1, \gamma_\lambda^2, \gamma_\lambda^w) = x_0^\top (K_0^1 + K_0^2) x_0$$

Using the relationship, specified in (4.8) and (4.9), between J^i and J_λ^i , the above inequality can be rewritten as

$$J^1(\gamma_\lambda^1, \gamma_\lambda^2, w) + J^2(\gamma_\lambda^1, \gamma_\lambda^2, w) \leq x_0^\top (K_0^1 + K_0^2) x_0 + \lambda \sum_{k=0}^{N-1} |w_k|^2 \quad (4.30)$$

By setting the initial value of system state $x_0 = 0$, we obtain

$$J^1(\gamma_\lambda^1, \gamma_\lambda^2, w) + J^2(\gamma_\lambda^1, \gamma_\lambda^2, w) \leq \lambda \sum_{k=0}^{N-1} |w_k|^2 \quad (4.31)$$

Then from (4.29), we have

$$\ll T_{\psi_\lambda} \gg \leq \sqrt{\lambda} \quad (4.32)$$

where $\psi_\lambda = (\gamma_\lambda^1, \gamma_\lambda^2)$.

Hence, given any $\lambda > 0$ satisfying condition (4.11) ($\Xi_k > 0$) of Theorem 8, the Nash strategies γ_λ^1 and γ_λ^2 , which depends on the value of λ , deliver the desired bound (4.32).

Remark 15 *In the above analysis, zero-valued initial system state is required to obtain the desired attenuation level.*

Definition 1 *Let $\Upsilon := \{\lambda > 0 \mid \Xi_k > 0\}$ where Ξ_k is defined in (4.11).*

Remark 16 Υ is not empty because the condition in (4.11) holds for λ sufficient large, and if a particular $\lambda \in \Upsilon$, then $(\lambda - \epsilon) \in \Upsilon$ for $\epsilon > 0$ sufficiently small.

Definition 2 $\hat{\lambda} := \inf \Upsilon$.

Remark 17 By Theorem 8, there does not exist any closed-loop Nash strategies which will deliver a bound in (4.32) that is smaller than $\sqrt{\hat{\lambda}}$. Hence

$$\sqrt{\lambda^*} = \sqrt{\hat{\lambda}}$$

where $\sqrt{\lambda^*}$ is defined in (4.28).

Theorem 9 Let $\hat{\lambda} \geq 0$ be a scalar specified in Definition 2, and $\psi_{\hat{\lambda}}^* = (\gamma_{\hat{\lambda}}^1, \gamma_{\hat{\lambda}}^2)$ be the state feedback Nash strategies defined in (4.15)-(4.16) with $\lambda = \hat{\lambda}$. Then, for finite-horizon LQSGs defined in (4.4)-(4.7), we have

$$\inf_{\psi \in \mathcal{M}_{CLPS}} \sup_{w \in \mathcal{H}_w} \frac{\|T_{\psi}(w)\|}{\|w\|} \equiv \inf_{\psi \in \mathcal{M}_{CLPS}} \ll T_{\psi} \gg = \ll T_{\psi_{\hat{\lambda}}^*} \gg = \sqrt{\hat{\lambda}}$$

i.e., the feedback Nash strategies $\psi_{\hat{\lambda}}^*$ solves the disturbance attenuation problem, and the optimal attenuation level equals $\sqrt{\hat{\lambda}}$.

Proof: The proof basically follows Theorem 8, Definition 2, and Remark 17. We provide here a formal proof for the **necessity** part of the result, that is the optimal attenuation level ($\inf_{\psi \in \mathcal{M}_{CLPS}} \ll T_{\psi} \gg$) can not be smaller than $\sqrt{\hat{\lambda}}$.

By contradiction, we suppose that the optimal attenuation level is $\tilde{\lambda} < \sqrt{\hat{\lambda}}$. Then, there exists $\tilde{\psi} = (\tilde{\gamma}^1, \tilde{\gamma}^2) \in \mathcal{M}_{CLPS}$ for some ϵ , $0 < \epsilon < \sqrt{\hat{\lambda}} - \sqrt{\tilde{\lambda}}$, such that

$$\ll T_{\tilde{\psi}} \gg = \sqrt{\tilde{\lambda}} + \epsilon := \sqrt{\tilde{\lambda}} < \sqrt{\hat{\lambda}}$$

From definition of $\ll T_{\psi} \gg$ in (4.29), we have

$$J^1(\tilde{\gamma}^1, \gamma^2, w) + J^2(\tilde{\gamma}^1, \gamma^2, w) - \tilde{\lambda} \sum_{k=0}^{N-1} |w_k|^2 = J_{\tilde{\lambda}}^1(\tilde{\gamma}^1, \gamma^2, w) + J_{\tilde{\lambda}}^2(\tilde{\gamma}^1, \gamma^2, w) \leq 0$$

for all possible w_k . However, it is impossible by Theorem 8 part (iii) because for $\forall \lambda < \hat{\lambda}$, either $J_{\hat{\lambda}}^1(\tilde{\gamma}^1, \gamma^2, w)$ or $J_{\hat{\lambda}}^2(\tilde{\gamma}^1, \gamma^2, w)$ will be unbounded from above. ■

4.3.4 Disturbance Attenuation Analysis of State Feedback Nash Strategies for Infinite-Horizon LQSGs

For the LQSG defined in (4.4) - (4.10), when $N \rightarrow +\infty$, it becomes an infinite-horizon LQSG problem. Let us examine the 2-player time-invariant cases defined as

$$x_{k+1} = Ax_k + B^1 u_k^1 + B^2 u_k^2 + Dw_k \quad (4.33)$$

with the constraints in (4.5). $x_k \in \mathcal{R}^n$, $u_k^i \in \mathcal{R}^{r_i}$, $1 \leq i \leq 2$, and $k = 0, 1, 2, \dots$.

The cost functions of the players are defined by

$$J_{\infty}^1 = \sum_{k=0}^{+\infty} (x_k^{\top} Q^1 x_k + u_k^{1\top} R^{11} u_k^1 + u_k^{2\top} R^{12} u_k^2) \quad (4.34)$$

$$J_{\infty}^2 = \sum_{k=0}^{+\infty} (x_k^{\top} Q^2 x_k + u_k^{1\top} R^{21} u_k^1 + u_k^{2\top} R^{22} u_k^2) \quad (4.35)$$

where all matrices are symmetric with $Q^i \geq 0$, (A, C^i) are observable (where $C^{i\top} C^i = Q^i$), and $R^{ij} > 0$ for $i = 1, 2$ and $j = 1, 2$. Then, the related zero-sum games have the following parameterized cost functions:

$$J_{\lambda, \infty}^1 = \sum_{k=0}^{+\infty} (x_k^{\top} Q^1 x_k + u_k^{1\top} R^{11} u_k^1 + u_k^{2\top} R^{12} u_k^2 - \delta_k^1 \lambda w_k^{\top} w_k) \quad (4.36)$$

$$J_{\lambda, \infty}^2 = \sum_{k=0}^{+\infty} (x_k^{\top} Q^2 x_k + u_k^{1\top} R^{21} u_k^1 + u_k^{2\top} R^{22} u_k^2 - \delta_k^2 \lambda w_k^{\top} w_k) \quad (4.37)$$

where $\lambda > 0$ and δ_k^i are defined in (4.10).

At each odd-valued time k , player 1 (u_k^1) will minimize its parameterized cost function $J_{\lambda, \infty}^1$ and disturbance (w_k) will try to maximize $J_{\lambda, \infty}^1$. On the other hand,

when k is even, $J_{\lambda,\infty}^2$ will be minimized by the control variable u_k^2 and maximized by the disturbance or uncertainty w_k .

Theorem 10 For K_k^i defined in (4.12)-(4.13) with $A_k = A$, $B_k^i = B^i$, $D_k = D$, $Q_k^i = Q^i$ and $R_k^{ij} = R^{ij}$, if Ξ_k in (4.11) is positive definite ($\Xi_k > 0$) for a fixed $\lambda > 0$, and there exists K_e^i and K_o^i , and for any fixed even-valued k , $K_k^i \rightarrow K_e^i$ and for any fixed odd-valued k , $K_k^i \rightarrow K_o^i$ as $N \rightarrow \infty$, then the state feedback Nash strategies of the game defined in (4.33)-(4.37) are given by

$$\gamma_k^1 = u_k^{1*} = \begin{cases} 0 & , \quad k \text{ is even} \\ -(R^{11})^{-1} B^{1\top} K_e^1 \Lambda_o^{-1} A x_k & , \quad k \text{ is odd} \end{cases} \quad (4.38)$$

$$\gamma_k^2 = u_k^{2*} = \begin{cases} -(R^{22})^{-1} B^{2\top} K_o^2 \Lambda_e^{-1} A x_k & , \quad k \text{ is even} \\ 0 & , \quad k \text{ is odd} \end{cases} \quad (4.39)$$

$$\gamma_k^w = w_k^* = \begin{cases} \frac{1}{\lambda} D^\top K_o^2 \Lambda_e^{-1} A x_k & , \quad k \text{ is even} \\ \frac{1}{\lambda} D^\top K_e^1 \Lambda_o^{-1} A x_k & , \quad k \text{ is odd} \end{cases} \quad (4.40)$$

where Λ_e and Λ_o are specified in (4.45) and (4.46), respectively. The optimal cost functions are

$$J_{\lambda,\infty}^{1*} = x_0^\top K_e^1 x_0$$

$$J_{\lambda,\infty}^{2*} = x_0^\top K_o^2 x_0$$

where x_0 is the initial state of the system defined in (4.33).

Proof: Using the conclusion and the same proof method of Theorem 8, we obtain the Nash strategies by taking the limits of the Nash solutions for the finite planning horizon cases as $N \rightarrow \infty$. ■

Remark 18 *From the conditions of Theorem 10, we obtain the coupled Discrete-time Algebraic Riccati Equations (DAREs) associated with LQSGs*

$$K_e^1 = Q^1 + A^\top [\Lambda_e^{-1}]^\top [K_o^1 + K_o^{2\top} B^2 (R^{22})^{-1} R^{12} (R^{22})^{-1} B^{2\top} K_o^2] \Lambda_e^{-1} A \quad (4.41)$$

$$K_o^1 = Q^1 + A^\top K_e^1 \Lambda_o^{-1} A \quad (4.42)$$

$$K_e^2 = Q^2 + A^\top K_o^2 \Lambda_e^{-1} A \quad (4.43)$$

$$K_o^2 = Q^2 + A^\top [\Lambda_o^{-1}]^\top [K_e^2 + K_e^{1\top} B^1 (R^{11})^{-1} R^{21} (R^{11})^{-1} B^{1\top} K_e^1] \Lambda_o^{-1} A \quad (4.44)$$

where

$$\Lambda_e = I + [B^2 (R^{22})^{-1} B^{2\top} - \frac{1}{\lambda} D D^\top] K_o^2 \quad (4.45)$$

$$\Lambda_o = I + [B^1 (R^{11})^{-1} B^{1\top} - \frac{1}{\lambda} D D^\top] K_e^1 \quad (4.46)$$

given $\lambda > 0$ satisfying

$$\Xi_e = \lambda I - D^\top K_o^2 D > 0 \quad (4.47)$$

$$\Xi_o = \lambda I - D^\top K_e^1 D > 0 \quad (4.48)$$

It is well known that there are no general existence conditions for the coupled DAREs associated with simultaneous games, see Başar and Olsder [37] (pp. 295-297). In this dissertation, we do not have the existence conditions for LQSGs either. We may attempt to obtain solutions by taking limits on K_k^i defined in (4.12)-(4.13), provided the limits exist. It is possible that there are other solutions, which also provide Nash strategies.

Remark 19 *The condition in Theorem 10 is equivalent to the condition that the associated coupled Riccati equations defined in (4.41)-(4.44) admit a positive semi-definite solution, which is known. Since in the infinite-horizon case ($N \rightarrow \infty$), we*

know that $x_N = 0$ due to the state feedback formulation. So we can choose arbitrary Q_N^1 and Q_N^2 . If the associated coupled Riccati equations admit a positive semi-definite solution (K_o^1, K_o^2) and (K_e^1, K_e^2) , we can choose $Q_N^i = K_o^i$ if N is odd-valued, or $Q_N^i = K_e^i$ if N is even-valued. Then the condition in Theorem 10 is satisfied. On the other hand, if the condition in Theorem 10 holds, (K_o^1, K_o^2) and (K_e^1, K_e^2) are a solution to the associated coupled Riccati equations.

Theorem 11 *In Theorem 10, if matrices $K_e^1, K_o^1, K_e^2, K_o^2, [B^1(R^{11})^{-1}B^{1\top} - \frac{1}{\lambda}DD^\top], [B^2(R^{22})^{-1}B^{2\top} - \frac{1}{\lambda}DD^\top]$, are positive definite, then the feedback system of LQSGs defined in (4.33)-(4.37) with the Nash strategies (4.38)-(4.40) is uniformly asymptotically stable.*

Proof: Define the time varying Lyapunov function as

$$V_k(x_k) = x_k^\top (\bar{K}_k^1 + \bar{K}_k^2) x_k \quad (4.49)$$

where

$$\bar{K}_k^i = \begin{cases} K_e^i & , \text{ if } k \text{ is even} \\ K_o^i & , \text{ if } k \text{ is odd} \end{cases} \quad (4.50)$$

By the conditions $K_e^i > 0$ and $K_o^i > 0$, we obtain

$$W_1(x_k) \leq V_k(x_k) \leq W_2(x_k) \quad (4.51)$$

where $W_1(x_k) = \min (x_k^\top (K_e^1 + K_e^2) x_k, x_k^\top (K_o^1 + K_o^2) x_k)$, which is positive definite.

$W_2(x_k) = x_k^\top (K_e^1 + K_e^2 + K_o^1 + K_o^2) x_k$, which is also positive definite.

Next, we will prove $V_{k+1}(x_{k+1}) - V_k(x_k) \leq 0$. Let us consider the odd-valued k first.

$$\begin{aligned}
V_{k+1}(x_{k+1}) - V_k(x_k) &= x_{k+1}^\top (K_e^1 + K_e^2) x_{k+1} - x_k^\top (K_o^1 + K_o^2) x_k \\
&= x_k^\top A^\top [\Lambda_o^{-1}]^\top (K_e^1 + K_e^2) \Lambda_o^{-1} A x_k - x_k^\top (K_o^1 + K_o^2) x_k \\
&= -x_k^\top Q^1 x_k - x_k^\top A^\top \left(I - [\Lambda_o^{-1}]^\top \right) K_e^1 \Lambda_o^{-1} A x_k \\
&\quad - x_k^\top [Q^2 + A^\top [\Lambda_o^{-1}]^\top K_e^1 B^1 (R^{11})^{-1} R^{21} (R^{11})^{-1} B^{1\top} K_e^1 \Lambda_o^{-1} A] x_k
\end{aligned}$$

In the above, the equality in the first line follows the definition of $V_k(x_k)$ in (4.49)-(4.50). The next equality comes from the result that $x_{k+1} = \Lambda_k^{-1} A x_k$ if γ_k^1, γ_k^1 , and γ_k^w specified in (4.38)-(4.40) are applied to the system in (4.33), which is justified in the proof of Theorem 8 part (i) and (ii). The equality in the third line follows the definition of K_o^1 in (4.42) and K_o^2 in (4.44).

Since $R^{21} > 0$, $Q^i \geq 0$, $R^{11} > 0$,

$$\left(I - [\Lambda_o^{-1}]^\top \right) K_e^1 \Lambda_o^{-1} = (\Lambda_o^{-1})^\top K_e^{1\top} \left[B^2 (R^{22})^{-1} B^{2\top} - \frac{1}{\lambda} D D^\top \right] K_e^1 \Lambda_o^{-1} \geq 0$$

we can obtain $V_{k+1}(x_{k+1}) - V_k(x_k) \leq 0$, which implies the uniform stability of the zero equilibrium.

Finally, an additional argument is needed to establish asymptotic stability. Since (A, C^1) and (A, C^2) are observable (where $C^{1\top} C^1 = Q^1$ and $C^{2\top} C^2 = Q^2$), it is not possible that $y_k^1(x_k) = C^1 x_k \equiv 0$ and $y_k^2(x_k) = C^2 x_k \equiv 0$ unless $x_k \equiv 0$. Then for some k ,

$$\begin{aligned}
V_{k+1}(x_{k+1}) - V_k(x_k) &= -x_k^\top Q^1 x_k - x_k^\top A^\top \left(I - [\Lambda_o^{-1}]^\top \right) K_e^1 \Lambda_o^{-1} A x_k \\
&\quad - x_k^\top [Q^2 + A^\top \Lambda_o^{-1} K_e^1 B^1 (R^{11})^{-1} R^{21} (R^{11})^{-1} B^{1\top} K_e^1 \Lambda_o^{-1} A] x_k \\
&\leq -(y_k^{1\top} y_k^1 + y_k^{2\top} y_k^2) < 0
\end{aligned}$$

Similarly, we can obtain the same conclusion for even-valued k . So the system is uniformly asymptotically stable. ■

Remark 20 *The closed-loop system is*

$$x_{k+1} = \Lambda_k^{-1} A x_k \quad (4.52)$$

if γ_k^1, γ_k^1 , and γ_k^w specified in (4.38)-(4.40) are applied to the system in (4.33), which is justified in the proof of Theorem 8 part (i) and (ii). From Theorem 11, we conclude that $\Lambda_k^{-1} A$ is Hurwitz, that is it has all its eigenvalues inside the unit circle.

Remark 21 *The conditions in Theorem 11, (that is, $K_e^1, K_o^1, K_e^2, K_o^2, [B^1(R^{11})^{-1}B^{1\top} - \frac{1}{\lambda}DD^\top]$, and $[B^2(R^{22})^{-1}B^{2\top} - \frac{1}{\lambda}DD^\top]$ are positive definite matrices), are sufficient but not necessary conditions of the conclusion that the feedback system of LQSGs defined in (4.33)-(4.37) with the Nash strategies (4.38)-(4.40) is uniformly asymptotically stable. The necessary and sufficient condition is that matrices $\Lambda_k^{-1} A$ is Hurwitz.*

Theorem 12 *For the same condition in Theorem 10, (that is, for K_k^i defined in (4.12)-(4.13) with $A_k = A$, $B_k^i = B^i$, $D_k = D$, $Q_k^i = Q^i$ and $R_k^{ij} = R^{ij}$, Ξ_k in (4.11) is positive definite ($\Xi_k > 0$) for a fixed $\lambda > 0$, and there exists K_e^i and K_o^i , and for any fixed even-valued k , $K_k^i \rightarrow K_e^i$ and for any fixed odd-valued k , $K_k^i \rightarrow K_o^i$ as $N \rightarrow \infty$), with the optimal state feedback Nash strategies specified in (4.38)-(4.39), the closed-loop system*

$$\begin{aligned} x_{k+1} &= A x_k + B^1 \gamma_k^1 + B^2 \gamma_k^2 + D w_k \\ &:= F_k x_k + D w_k \end{aligned} \quad (4.53)$$

where

$$F_k := \begin{cases} F_e := [I - B^2(R^{22})^{-1}B^{2\top}K_o^2\Lambda_e^{-1}]A & , \quad k \text{ is even} \\ F_o := [I - B^1(R^{11})^{-1}B^{1\top}K_e^1\Lambda_o^{-1}]A & , \quad k \text{ is odd} \end{cases} \quad (4.54)$$

is bounded input and bounded output stable with respect to the system disturbances w_k .

Proof: We only need to prove that F_k is *Hurwitz*. From Theorem 10, we know that the optimal cost functions are

$$J_{\lambda,\infty}^{1*} = x_0^\top K_e^1 x_0$$

$$J_{\lambda,\infty}^{2*} = x_0^\top K_e^2 x_0$$

where x_0 is the initial state of the system defined in (4.33). From the proof of Theorem 8, we know, for any noises or disturbances w_k ,

$$J_{\lambda,\infty}^1(\gamma_{[0,\infty]}^1, \gamma_{[0,\infty]}^2, w_{[0,\infty]}) \leq J_{\lambda,\infty}^1(\gamma_{[0,\infty]}^1, \gamma_{[0,\infty]}^2, \gamma_{[0,\infty]}^w) = J_{\lambda,\infty}^{1*} = x_0^\top K_e^1 x_0$$

$$J_{\lambda,\infty}^2(\gamma_{[0,\infty]}^1, \gamma_{[0,\infty]}^2, w_{[0,\infty]}) \leq J_{\lambda,\infty}^2(\gamma_{[0,\infty]}^1, \gamma_{[0,\infty]}^2, \gamma_{[0,\infty]}^w) = J_{\lambda,\infty}^{2*} = x_0^\top K_e^2 x_0$$

where $\gamma_{[0,\infty]}^1$, $\gamma_{[0,\infty]}^2$, $\gamma_{[0,\infty]}^w$, and $w_{[0,\infty]}$ are the action sets for player 1, player 2, disturbances with strategies γ_k^1 , γ_k^2 , γ_k^w (specified in (4.15)-(4.17)) and any disturbance strategy generating w_k , respectively. So by setting $w_k \equiv 0$, we obtain bounded J_∞^1 and J_∞^2 with state feedback Nash strategies γ_k^1 and γ_k^2 .

Then we will use an argument, which is similar to that used in the linear regulator theory, to prove that $x_k \rightarrow 0$. From the boundedness of the upper values of J_∞^1 and J_∞^2 , we can conclude that

$$x_k^\top Q^1 x_k + \gamma_k^{1\top} R^{11} \gamma_k^1 + \gamma_k^{2\top} R^{12} \gamma_k^2 \rightarrow 0 \quad \text{as} \quad k \rightarrow +\infty$$

$$x_k^\top Q^2 x_k + \gamma_k^{2\top} R^{22} \gamma_k^2 + \gamma_k^{1\top} R^{21} \gamma_k^1 \rightarrow 0 \quad \text{as} \quad k \rightarrow +\infty$$

We denote $B^1 \gamma_k^1 + B^2 \gamma_k^2 = s_k x_k$, then $x_{k+1} = Ax_k + s_k x_k$. When $k \rightarrow +\infty$,

$$C^1 x_k \rightarrow 0, \quad C^2 x_k \rightarrow 0, \quad \text{and} \quad s_k x_k \rightarrow 0$$

because $Q^1 = C^{1\top}C^1$, $Q^2 = C^{2\top}C^2$, $R^{ij} > 0$ and $(\gamma_k^i \rightarrow 0) \Rightarrow (s_k x_k \rightarrow 0)$. Then for the next time step $k+1$, we have

$$C^1 x_{k+1} \rightarrow 0, \quad C^2 x_{k+1} \rightarrow 0, \quad \text{and} \quad s_{k+1} x_{k+1} \rightarrow 0$$

Since $x_{k+1} = Ax_k + s_k x_k$ and $s_k x_k \rightarrow 0$,

$$C^1 x_{k+1} \rightarrow 0 \quad \Rightarrow \quad C^i A x_k \rightarrow 0$$

Similarly, we have

$$C^1 x_{k+2} \rightarrow 0, \quad C^2 x_{k+2} \rightarrow 0, \quad \text{and} \quad s_{k+2} x_{k+2} \rightarrow 0 \quad \Rightarrow \quad C^i A^2 x_k \rightarrow 0$$

$$\dots \quad \dots \quad \dots$$

$$C^1 x_{k+n-1} \rightarrow 0, \quad C^2 x_{k+n-1} \rightarrow 0, \quad \text{and} \quad s_{k+n-1} x_{k+n-1} \rightarrow 0 \quad \Rightarrow \quad C^i A^{n-1} x_k \rightarrow 0$$

By the observability conditions (that is, (A, C^i) are observable (where $C^{i\top}C^i = Q^i$)), $C^1 A^i x_k \rightarrow 0$ or $C^2 A^i x_k \rightarrow 0$, $i = 0, 1, \dots, n-1$ implies that $x_k \rightarrow 0$. Hence F_k is *Hurwitz*. This completes our proof. ■

Definition 3 For the system defined in (4.33)-(4.37), we define

$$\Upsilon_\infty := \{\lambda_\infty > 0 \mid \text{for } \forall \lambda > \lambda_\infty, \exists \Xi_e > 0 \text{ and } \Xi_o > 0\}$$

where Ξ_e and Ξ_o are defined in (4.47) and (4.48), respectively.

Remark 22 Υ_∞ is not empty because for λ sufficient large, $\Xi_e > 0$ and $\Xi_o > 0$.

Definition 4 $\hat{\lambda}_\infty := \inf \Upsilon_\infty$.

Remark 23 From Definition 4 and Theorem 12, we know that $\forall \lambda \geq \hat{\lambda}_\infty$, the closed-loop system with optimal state feedback Nash strategies specified in (4.38)-(4.39) is bounded input and bounded output (BIBO) stable with respect to the system disturbances w_k .

Theorem 13 Let $\hat{\lambda}_\infty \geq 0$ be a scalar specified in Definition 4, and $\psi_{\hat{\lambda}_\infty}^* = (\gamma_{\hat{\lambda}_\infty}^1, \gamma_{\hat{\lambda}_\infty}^2)$ be the state feedback Nash strategies defined in (4.38)-(4.39) with $\lambda = \hat{\lambda}_\infty$. Then, for infinite-horizon LQSGs defined in (4.33)-(4.37), we have

$$\inf_{\psi \in \mathcal{M}_{CLPS}^\infty} \sup_{w \in \mathcal{H}_w} \frac{\|T_\psi(w)\|}{\|w\|} \equiv \inf_{\psi \in \mathcal{M}_{CLPS}^\infty} \ll T_\psi \gg = \sqrt{\hat{\lambda}_\infty}$$

i.e., that the feedback Nash strategies $\psi_{\hat{\lambda}_\infty}^*$ solve the disturbance attenuation problem, and the optimal attenuation level equals $\sqrt{\hat{\lambda}_\infty}$.

Furthermore, the controllers $\psi_{\hat{\lambda}_\infty}^* = (\gamma_{\hat{\lambda}_\infty}^1, \gamma_{\hat{\lambda}_\infty}^2)$ lead to a bounded input bounded output stable system.

Proof: The first part of the solution for the disturbance attenuation problem can be proved by a similar method used in Theorem 9. The second part for bounded input bounded output stability follows Theorem 12. ■

4.4 Simulations

A numerical example of second-order system is given in this section to illustrate our disturbance attenuation oriented state feedback Nash strategies of LQSGs with disturbance. Consider the system defined in (4.33) - (4.37) with

$$A = \begin{bmatrix} 0.5 & 0.0 \\ 0.0 & 0.6 \end{bmatrix}, Q_1 = \begin{bmatrix} 1.1 & 0.0 \\ 0.0 & 1.0 \end{bmatrix}, Q_2 = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.1 \end{bmatrix}$$

$B^1 = [0, 1]^\top$, $B^2 = [1, 0]^\top$, $R_{11} = 1.1$, $R_{12} = 2.0$, $R_{21} = 2.5$, $R_{22} = 1.5$, and $D = [1, 1]^\top$. We first calculate $\hat{\lambda}_\infty$ in Definition 4 and the associated K_k^1 , K_k^2 in Theorem 10. By numerical methods, we obtain

$$\hat{\lambda}_\infty = 6.2928$$

With the optimal $\hat{\lambda}_\infty$, Ξ_k is shown in Fig. 4.1, K_k^1 in Fig. 4.2, K_k^2 in Fig. 4.4, Λ_k in Fig. 4.6. When $N \rightarrow \infty$, we found that K_o^i , K_e^i , Λ_o and Λ_e exist.

$$\begin{aligned} K_o^1 &= \begin{bmatrix} 1.6925 & 0.2090 \\ 0.2090 & 1.3960 \end{bmatrix}, K_e^1 = \begin{bmatrix} 1.6692 & 0.7708 \\ 0.7708 & 3.0820 \end{bmatrix} \\ K_o^2 &= \begin{bmatrix} 1.9575 & 0.6536 \\ 0.6536 & 2.1327 \end{bmatrix}, K_e^2 = \begin{bmatrix} 1.3242 & 0.3568 \\ 0.3568 & 2.4832 \end{bmatrix} \\ \Lambda_o &= \begin{bmatrix} 0.6123 & -.6122 \\ 0.3130 & 3.1895 \end{bmatrix}, \Lambda_e = \begin{bmatrix} 1.8901 & -.0070 \\ -.4149 & 0.5572 \end{bmatrix} \end{aligned}$$

The above matrices are positive definite since their eigenvalues, which are shown in Fig. 4.3, Fig. 4.5 and Fig. 4.7, are positive.

$$\begin{aligned} \text{eig}(K_o^1) &= \begin{bmatrix} 1.8005 \\ 1.2880 \end{bmatrix}, \text{eig}(K_e^1) = \begin{bmatrix} 1.3300 \\ 3.4211 \end{bmatrix}, \text{eig}(K_o^2) = \begin{bmatrix} 1.3856 \\ 2.7045 \end{bmatrix} \\ \text{eig}(K_e^2) &= \begin{bmatrix} 1.2232 \\ 2.5842 \end{bmatrix}, \text{eig}(\Lambda_o) = \begin{bmatrix} 0.6889 \\ 3.1129 \end{bmatrix}, \text{eig}(\Lambda_e) = \begin{bmatrix} 1.8922 \\ 0.5550 \end{bmatrix} \end{aligned}$$

and $\Xi_o = \hat{\lambda}_\infty - D^\top K_e^1 D = 0.0001 > 0$, $\Xi_e = \hat{\lambda}_\infty - D^\top K_o^2 D = 0.8955 > 0$.

As shown in Fig. 4.8 and Fig. 4.9, the closed-loop system with state feedback Nash Strategies for real decision makers (u_k^1 , u_k^2) and disturbance (w_k) are asymptotically stable.

When $w_k \sim \mathcal{N}(0, 0.1)$, the feedback Nash strategies for real decision makers (u_k^1 , u_k^2) are shown in Fig. 4.10. The closed-loop system is BIBO stable, as shown in Fig. 4.11. The cost-to-go V_k^1 and V_k^2 , as shown in Fig. 4.12, are unbounded from above

since x_k does not converge to 0 due to the system disturbance w_k . However, as shown in Fig. 4.13, $V_1^1 + V_1^2$ is bounded from above by V_1^w , where $V_k^w = \hat{\lambda}_\infty \sum_{i=k}^{N-1} w_i^\top w_i$. From (4.30), we know that for $\forall w_k$, $(V_1^1) + (V_1^2) \leq V_1^w$.

4.5 Chapter Summary

In this chapter, the disturbance attenuation analysis of state feedback Nash strategies was considered based on a H_∞ -optimal control problem, which was converted into a zero-sum game. In this approach, the uncertain parameters were controlled by a virtual player which was maximizing the cost function that was being minimized by the real decision makers. For finite-horizon LQSGs, we derived state feedback Nash strategies with the optimal attenuation level $\sqrt{\hat{\lambda}} > 0$. We extended the approach to infinite-horizon LQSGs. A lower bound $\hat{\lambda}_\infty$ was provided in the sense that for every $\lambda < \hat{\lambda}_\infty$ the values of the LQSGs are unbounded. We also proved that the feedback system with state feedback Nash strategies was Bounded Input Bounded Output (BIBO) stable with respect to the disturbances.

Figure 4.1: Ξ_k , when $\hat{\lambda}_\infty = 6.2928$. $(\Xi_k) \rightarrow \Xi_o = 0.0001 > 0$ for odd-valued k , and $(\Xi_k) \rightarrow \Xi_e = 0.8955 > 0$ for even-valued k

Figure 4.2: K_k^1 , where $(K_k^1)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^1 , which is obtained from (4.12). K_k^1 converges.

Figure 4.3: Eigenvalues of K_k^1 . K_o^1 and K_e^1 exist, and both are positive definite.

Figure 4.4: K_k^2 , where $(K_k^2)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix K_k^2 , which is obtained from (4.13). K_k^2 converges.

Figure 4.5: Eigenvalues of K_k^2 . K_o^2 and K_e^2 exist, and both are positive definite.

Figure 4.6: Λ_k , where $(\Lambda_k)_{ij}$ is the matrix element at i^{th} row and j^{th} column of matrix Λ_k , which is obtained from (4.14). Λ_k converges.

Figure 4.7: Eigenvalues of Λ_k . Λ_o and Λ_e exist, and both are positive definite.

Figure 4.8: State feedback Nash strategies for real decision makers (u_k^1, u_k^2) and disturbance (w_k) .

Figure 4.9: Asymptotically stable closed-loop system with state feedback Nash Strategies for real decision makers and disturbance.

Figure 4.10: State feedback Nash strategies for real decision makers (u_k^1, u_k^2) . w_k is white noise with $\mu_w = 0$ and $\sigma_w^2 = 0.1$

Figure 4.11: The closed-loop system, with state feedback Nash strategies for real decision makers (u_k^1, u_k^2) is BIBO stable for white noise $w_k \sim \mathcal{N}(0, 0.1)$.

Figure 4.12: Cost-to-Go V_k^1 and V_k^2 when $w_k \sim \mathcal{N}(0, 0.1)$. Note that $J^{i*} = V_1^i$.

Figure 4.13: When $w_k \sim \mathcal{N}(0, 0.1)$, $(V_1^1) + (V_1^2) < V_1^w$, where $V_k^w = \hat{\lambda}_\infty \sum_{i=k}^{N-1} w_i^\top w_i$. From (4.30), we know that for $\forall w_k$, $(V_1^1) + (V_1^2) \leq V_1^w$

CHAPTER 5

CONCLUSIONS

Sequential games are natural and useful frameworks to address some real problems, such as the “Action-Reaction-Counteraction” [41, 42] paradigm used in military intelligence and advertising campaigns strategies of several competing firms in economics. They are conventionally represented by game trees (the extensive form) and solved using the concept of rollback, or sub-game perfect equilibrium. These current tools lack the sophistication to study games in which the players have relatively large, or even infinitely large, planning horizons. The theoretical results for dynamic non-cooperative sequential games are very limited. This dissertation research focused on existence, stability analysis, adaptation design, and disturbance attenuation analysis of state feedback Nash strategies for LQSGs from a system control perspective. We now review the major findings/contributions of this work and propose ideas for future work.

5.1 Contributions

5.1.1 Existence and Stability Analysis

In Chapter 2, using Dynamic Programming and Isaacs’ equation [46], we derived state feedback Nash strategies for Linear Quadratic Sequential Games (LQSGs) with

finite planning horizons. We showed that the existence of unique state feedback Nash strategies were guaranteed, , under some regularity assumptions on the weighting matrices. For the same system, we provided open-loop Nash strategies via a Hamiltonian based method. We also extended the result to the M -player case. Given that the coupled algebraic Riccati equations have a positive definite solution, there exists state feedback Nash strategies for the general infinite-horizon LQSGs. In particular, for infinite-horizon scalar LQSGs, we proved that the normal regularity conditions, such as (A, B^i) is stabilizable and (A, C^i) is detectable (where $C^{i\top} C^i = Q^i$), guaranteed a positive definite solution to the associated coupled algebraic Riccati equations.

Based on Lyapunov's method for time-variant systems, we proved uniform asymptotic stability of close-loop LQSGs with state feedback Nash strategies under the assumption that the discrete algebraic Riccati equations have a positive definite solution.

5.1.2 Adaptation Design and Convergence Analysis

In Chapter 3, we combined Fictitious Play (FP) [6, 7], (in which each player presumes that its opponents are playing Nash strategies, then a best response to those beliefs is chosen as a strategy), and conventional adaptive control techniques such as normalized gradient algorithm and Recursive Least Square (RLS) to design adaptation schemes for decision makers in linear quadratic games with state feedback Nash strategies (we assume that the players' controls were scalars).

We first considered the adaptivity issues for simultaneous linear quadratic games in a one-side adaptation scheme, in which only one of the two players has perfect information of the cost functions of both players, i.e., one player knows the exact Q^1 ,

$Q^2, R^{11}, R^{12}, R^{21}, R^{22}$, while the other player only has access to its cost function and does not know the parameters of the cost function of its opponent. Following the concept of FP, we assumed that both players will apply the Nash strategies based on the available information set. That is the player with the perfect information set will play real state feedback Nash strategies while the player without perfect information knows in advance that its opponent will implement state feedback Nash strategies, and it can measure the state perfectly, based on which it can design his adaptive control. We designed an indirect adaptive method to estimate the control gain of the player with perfect information. Assuming that stabilizability and detectability conditions are satisfied, we proved that the player with imperfect information can always calculate its best-response control gain, based on its estimate of the other player's control gain substituted in the coupled discrete-time algebraic Riccati equations. Convergence was proved and the persistent excitation conditions were satisfied by two methods: reference signal tracking and small system noises. We extended our approach to LQSGs via specially designed past system state measurement ϕ_k .

In the second part of this chapter, adaptivity analysis was considered in a more useful and practical two-side adaptation case, in which we assumed that each player has access to its own cost function as well as the system dynamics, and does not know the parameters of the cost function of its opponent. i.e., player 1 knows $Q^1, R^{12}, R^{11}, A, B^1$ and B^2 while player 2 has access to $Q^2, R^{21}, R^{22}, A, B^1$ and B^2 . So both players need to estimate the control gain of its opponent under a FP framework, in which we assumed that each player knows in advance that its opponent will implement state feedback Nash strategies. Then, player 2 (or player 1) will estimate the control gain L^1 (or L^2) first, and then calculate/estimate its own best response control gain L^2 (or L^1).

(or L^1) from the related coupled algebraic discrete-time Riccati equations. Existence was guaranteed via a transformed regular discrete-time algebraic Riccati equation. We provided adaptation design based on normalized gradient algorithm [58]. With the help of two properties of “reaction-curve” like relations between control gains of two players, we proved that the parameter (control gain) estimation convergence given that persistent excitation conditions [45] were satisfied. Similar to the one-side adaptation scheme, we extended the adaptive law and convergence proof to LQSGs.

5.1.3 Disturbance Attenuation Analysis

In Chapter 4, we extended a zero-sum game theoretical approach [38] for H_∞ -optimal control problems to address the disturbance attenuation issues of the state feedback Nash strategies for LQSGs with uncertainties or disturbances. In the approach, the disturbance attenuation analysis of state feedback Nash strategies was considered based on a H_∞ -optimal control problem, which was converted into a zero-sum game, in which the uncertain parameters were controlled by a virtual player which was maximizing the cost functions that were being minimized by the real decision makers. That is, at each odd-valued time k , player 1 (u_k^1) will minimize its parameterized cost function J_λ^1 and disturbance (w_k) will try to maximize J_λ^1 ; and when k is even, J_λ^2 will be minimized by the control variable u_k^2 and maximized by the disturbance or uncertainty w_k .

For finite-horizon LQSGs, we solved a series of zero-sum games. We first provided, for finite-horizon LQSGs, state feedback Nash strategies with the optimal attenuation level $\sqrt{\hat{\lambda}} > 0$. We extended the approach to infinite-horizon LQSGs. A lower bound $\hat{\lambda}_\infty$ was provided in the sense that for every $\lambda < \hat{\lambda}_\infty$ the values of the LQSGs are

unbounded. We also proved that the feedback system with state feedback Nash strategies was Bounded Input Bounded Output (BIBO) stable for the disturbances.

5.2 Future Directions

There are several problems that remain open in the research in this dissertation. Clearly, it would be useful to consider the feedback Nash strategies for LQSGs under different information structures [37], such as one-step delay ($u_k = f(x_{k-1})$) and noisy state measurement ($u_k = f(x_k + w_k)$) rather than the perfect state ($u_k = f(x_k)$) one used in this dissertation.

The associated discrete-time matrix Riccati equations are very useful for analyzing the stability and disturbance attenuation analysis of sequential games. Existence conditions for continuous time Riccati equations can be found in [59]. However, for the nonzero-sum game case, the related equations are coupled. So far, there are no general existence conditions for this kind of Riccati equations. For our case of sequential games, the problem is more complex because of the nonlinearity from the sequential plays. Another future research direction is to find existence conditions for bounded solutions of general coupled discrete-time algebraic Riccati equations associated with LQSGs. In this dissertation, we have solved the scalar case.

In the two-side adaptation scheme in Chapter 3, to satisfy excitation conditions [45], we proposed two methods: reference signal tracking and small system noise. We only proved convergence for reference signal tracking case. It is possible to prove the unbiasedness of the adaptive law specified in Remark 12 of Chapter 3.

There is another future direction in the two-side adaptation scheme investigated in Chapter 3. We conjectured that similar results such as convergence and existence

are achievable for linear quadratic games with vector control inputs for each player. In this general case, regression models are needed to design adaptive laws. Then the estimated parameters θ^i is a function of the control gain L^i . If the relation is unique, then we can obtain $\hat{L}^i$ from the estimated parameters $\hat{\theta}^i$, that is, $\hat{L}^i = f^i(\hat{\theta}^i)$. If the function f^i can preserve the relation in (3.41)-(3.42), then convergence is guaranteed.

In summary, this dissertation has presented stability analysis, adaptation design, and disturbance attenuation of LQSGs with state feedback Nash strategies. Future research needs to broaden the results in the directions of more information structure, practical usage, and fewer assumptions.

BIBLIOGRAPHY

- [1] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.
- [2] J. F. Nash, Jr., “Equilibrium points in N -person games,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 36, pp. 48–49, 1950.
- [3] J. F. Nash, Jr., “Non-cooperative games,” *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.
- [4] J. F. Nash, Jr., “The bargaining problem,” *Econometrica*, vol. 188, pp. 155–162, 1950.
- [5] J. F. Nash, Jr., “Two person cooperative games,” *Econometrica*, vol. 21, pp. 128–140, 1953.
- [6] G. W. Brown, “Iterative solutions of games by fictitious play,” in *Activity Analysis of Production and Allocation* (T. C. Koopmans, ed.), New York: Wiley, 1951.
- [7] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge: MIT Press, 1998.
- [8] D. Gillies, “Solutions to general non-zero-sum games,” in *Contributions to the Theory of Games* (A. W. Tucker and R. D. Luce, eds.), vol. IV (Annals of Mathematics Studies, 40), pp. 47–85, Princeton: Princeton University Press, 1959.
- [9] L. S. Shapley, “A value for n -person games,” in *Contributions to the Theory of Games* (A. W. Tucker and R. D. Luce, eds.), vol. II (Annals of Mathematics Studies, 28), pp. 307–317, Princeton: Princeton University Press, 1953.
- [10] L. S. Shapley, “Stochastic games,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 39, pp. 1095–1100, 1953.

- [11] H. W. Kuhn, "Extensive games and the problem of information," in *Contributions to the Theory of Games* (A. W. Tucker and R. D. Luce, eds.), vol. II (Annals of Mathematics Studies, 28), pp. 193–216, Princeton: Princeton University Press, 1953.
- [12] R. J. Aumann and M. Maschler, "Game - theoretic aspects of gradual disarmament," in *Report to the U.S. Arms Control and Disarmament Agency ST-80*, ch. V, Princeton: Mathematica, 1966.
- [13] R. J. Aumann, "Survey of repeated games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pp. 11–42, Zurich: Bibliographisches Institut, 1981.
- [14] A. Neyman, "Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma," *Economic Letters*, vol. 19, pp. 227–229, 1985.
- [15] A. Rubinstein, "Finite automata play the repeated prisoner's dilemma," *Journal of Economic Theory*, vol. 39, pp. 83–96, 1986.
- [16] J. M. Smith, "Game theory and the evolution of fighting," in *On Evolution*, pp. 8–28, Edinburgh: Edinburgh University Press, 1972.
- [17] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, pp. 15–18, 1973.
- [18] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, vol. 50, no. 1, pp. 67–96, 1974.
- [19] R. Selten, "Re-examination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, vol. 4, pp. 25–55, 1975.
- [20] D. M. Kreps and R. Wilson, "Sequential equilibria," *Econometrica*, vol. 50, no. 4, pp. 863–894, July 1982.
- [21] V. P. Crawford, "Equilibrium without independence," *Journal of Economic Theory*, vol. 50, no. 1 (Feb.), pp. 127–154, 1990.
- [22] P. Reny, "On the existence of pure and mixed strategy Nash equilibria in discontinuous games," *Econometrica*, vol. 67, no. 5 (Sep.), pp. 1029–1056, 1999.
- [23] S. Pápai, "Unique stability in simple coalition formation games," *Games and Economic Behavior*, vol. 48, pp. 337–354, 2004.
- [24] A. Heifetz and E. Segev, "The evolutionary role of toughness in bargaining," *Games and Economic Behavior*, vol. 49, pp. 117–134, 2004.

- [25] Jérôme Renault and Tristan Tomala, “Communication equilibrium payoffs in repeated games with imperfect monitoring,” *Games and Economic Behavior*, vol. 49, pp. 313–344, 2004.
- [26] D. Samet and Z. Safra, “A family of ordinal solutions to bargaining problems with many players,” *Games and Economic Behavior*, vol. 50, pp. 89–106, 2005.
- [27] M. B. Stinchcombe, “Nash equilibrium and generalized integration for infinite normal form games,” *Games and Economic Behavior*, vol. 50, pp. 332–365, 2005.
- [28] J. von Neumann, “On the theory of games of strategy,” in *Contributions to the Theory of Games* (A. W. Tucker and R. D. Luce, eds.), vol. IV (Annals of Mathematics Studies, 40), pp. 13–42, Princeton: Princeton University Press, 1959.
- [29] A. Ambrosetti and P. H. Rabinowitz, “Dual variational methods in critical point theory and applications,” *J. Funct. Anal.*, vol. 14, pp. 349–381, 1973.
- [30] R. P. Isaacs, *Differential games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. New York: John Wiley & Sons, 1965.
- [31] H. von Stackelberg, *The theory of the market Economy*. London, U.K.: Oxford Univ. Press, 1952. Transl: A.T. Peacock.
- [32] C. I. Chen and J. B. Cruz, Jr., “Stackelberg solution for two-person games with biased information patterns,” *IEEE Trans. Automat. Contr.*, vol. AC-17, no. 6, pp. 791–798, 1972.
- [33] M. Simaan and J. B. Cruz, Jr., “On the Stackelberg strategy in nonzero-sum games,” *Journal of Optimization Theory and Applications*, vol. 11, no. 5, pp. 533–555, 1973.
- [34] M. Simaan and J. B. Cruz, Jr., “Additional aspects of the Stackelberg strategy in nonzero-sum games,” *Journal of Optimization Theory and Applications*, vol. 11, no. 6, pp. 613–626, 1973.
- [35] M. Li, J. B. Cruz, Jr., and M. A. Simaan, “An approach to discrete-time incentive feedback Stackelberg games,” *IEEE Trans. Syst., Man, Cybern. A*, vol. 32, no. 4, pp. 472 – 481, July 2002.
- [36] T. Başar and H. Selbuz, “Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems,” *IEEE Trans. Automat. Contr.*, vol. AC-24, no. 2, pp. 166–179, 1979.

- [37] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, second ed., January, 1999.
- [38] T. Başar and P. Bernhard, *H_∞ -Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Boston, MA: Birkhäuser, 1991.
- [39] P. Stone and M. Veloso, “Multiagent systems: A survey from a machine learning perspective,” in *Autonomous Robots*, vol. 8, pp. 345–383, Netherlands: Springer, June 2000.
- [40] C. J. C. H. Watkins, “Learning from delayed rewards,” *Ph.D Thesis, Cambridge University, Cambridge, England*, 1989.
- [41] M. Horrell, *Action, Reaction and Counter-action; A Brief Review of Non-white Political Movements in South Africa*. Johannesburg: South African Institute of Race Relations, 1971.
- [42] C. J. Ancker, III and M. D. Burke, “Doctrine for asymmetric warfare,” *Military Review*, July-August 2003.
- [43] W. White, “Representation and approximation of noncooperative sequential games,” *SIAM Journal on Control and Optimization*, vol. 18, pp. 33–48, 1980.
- [44] M. Breton, A. Alj, and A. Haurie, “Sequential Stackelberg equilibria in two-person games,” *Journal of Optimization Theory and Applications*, vol. 59, no. 1, pp. 71–97, Oct. 1988.
- [45] K. J. Åström and B. Wittenmark, *Adaptive Control*. Addison-Wesley Series in Electrical Engineering: Control Engineering, Addison-Wesley, second ed., 1995.
- [46] R. Isaacs, *Differential Games*. Huntington, New York: Kruger Publishing Company, 1975.
- [47] J. P. Hespanha, “Uniform stability of switched linear systems: extensions of LaSalle’s Invariance Principle,” *IEEE Trans. Automat. Contr.*, vol. 49, no. 4, pp. 470–482, Apr. 2004.
- [48] J. P. Hespanha, D. Liberzon, D. Angeli, and E. D. Sontag, “Nonlinear norm-observability notions and stability of switched systems,” *IEEE Trans. Automat. Contr.*, vol. 50, no. 2, pp. 154–168, Feb. 2005.
- [49] J. Alster and P. R. Belanger, “A technique of dual adaptive control,” *Automatica*, vol. 10, pp. 627–634, 1974.
- [50] Y. Bar-Shalom and E. Tse, “Dual effect, certainty equivalence and separation in stochastic control,” *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 494–500, 1974.

- [51] B. B. Peterson and K. S. Narendra, "Bounded error adaptive control," *IEEE Trans. Automat. Contr.*, vol. 27, no. 6, pp. 1161–1168, 1982.
- [52] G. Kreisselmeier and B. Anderson, "Robust model reference adaptive control," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 127–133, 1986.
- [53] A. S. Morse, "Global stability of parameter-adaptive control systems," *IEEE Trans. Automat. Contr.*, vol. 25, pp. 433–439, 1980.
- [54] K. S. Narendra, Y. H. Lin, and L. S. Valavani, "Stable adaptive controller design, part ii: Proof of stability," *IEEE Trans. Automat. Contr.*, vol. 25, pp. 440–448, 1980.
- [55] G. C. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete time multivariable control," *IEEE Trans. Automat. Contr.*, vol. 25, pp. 449–456, 1980.
- [56] K. J. Åström and B. Wittenmark, "On self-tuning regulators," *Automatica*, vol. 9, pp. 185–199, 1973.
- [57] J. M. Smith, *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- [58] G. Tao, *Adaptive Control Design and Analysis*. Adaptive and Learning Systems for Signal Processing, Communications and Control Series, Hoboken, N.J.: Wiley-Interscience, 2003.
- [59] W. M. Wonham, "On a matrix Riccati equation of stochastic control," *SIAM Journal on Control*, vol. 6, no. 4, pp. 681–697, 1968.
- [60] H. S. Black, "Stabilized feedback amplifiers," *Bell Syst. Tech. J.*, vol. 13, pp. 1–18, 1943.
- [61] H. Nyquist, "Regeneration theory," *Bell Syst. Tech. J.*, vol. 11, pp. 126–147, 1932.
- [62] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. Princeton, NJ: Van Nostrand, 1945.
- [63] I. M. Horowitz, *Synthesis of Feedback Systems*. New York: Academic Press, 1963.
- [64] J. B. Cruz, Jr. and W. R. Perkins, "A new approach to the sensitivity problem in multivariable feedback system design," *IEEE Trans. Automat. Contr.*, vol. AC-9, pp. 216–223, July 1964.
- [65] J. B. Cruz, Jr., J. S. Freudenberg, and D. P. Looze, "A relationship between sensitivity and stability of multivariable feedback systems," *IEEE Trans. Automat. Contr.*, vol. 26, pp. 66 – 74, Feb. 1981.

- [66] G. Zames and B. A. Francis, "Feedback, minimax sensitivity, and optimal robustness," *IEEE Trans. Automat. Contr.*, vol. 28, pp. 585–601, May 1983.
- [67] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Automat. Contr.*, vol. 26, pp. 301–320, Apr. 1981.
- [68] G. Zames and D. Bensoussan, "Multivariable feedback, sensitivity, and decentralized control," *IEEE Trans. Automat. Contr.*, vol. 28, pp. 1030–1035, Nov. 1983.
- [69] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 831–847, Aug. 1989.
- [70] A. A. Stoorvogel, A. Saberi, and B. M. Chen, "The discrete-time H_∞ control problem with strictly proper measurement feedback," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1936–1939, Sept. 1994.
- [71] R. K. Ragade and I. G. Sarma, "A game theoretic approach to optimal control in the presence of uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-12, no. 4, pp. 395–401, Aug. 1967.
- [72] P. Dorato and A. Kestenbaum, "Application of game theory to the sensitivity design of optimal systems," *IEEE Trans. Automat. Contr.*, vol. AC-12, no. 1, pp. 85–87, Feb. 1967.
- [73] D. Salmon, "Minimax controller design," *IEEE Trans. Automat. Contr.*, vol. AC-13, no. 4, pp. 369–376, Aug. 1968.
- [74] H. S. Witsenhausen, "A minimax control problem for sampled linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-13, no. 1, pp. 5–21, Feb. 1968.
- [75] C. R. Johnson, "Positive definite matrices," *Amer. Math. Monthly*, vol. 77, pp. 259–264, 1970.