STARS

University of Central Florida
STARS

Electronic Theses and Dissertations, 2004-2019

2013

Differential Games For Multi-agent Systems Under Distributed Information

Wei Lin
University of Central Florida

Part of the Electrical and Electronics Commons

Find similar works at: https://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

University of Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

DIFFERENTIAL GAMES FOR MULTI-AGENT SYSTEMS UNDER DISTRIBUTED
INFORMATION

by

WEI LIN
Bachelor of Engineering, Tongji University, 2007
Master of Science, Boston University, 2009

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical Engineering and Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2013

Major Professor: Zhihua Qu

ABSTRACT

In this dissertation, we consider differential games for multi-agent systems under distributed information where every agent is only able to acquire information about the others according to a directed information graph of local communication/sensor networks. Such games arise naturally from many applications including mobile robot coordination, power system optimization, multi-player pursuit-evasion games, etc. Since the admissible strategy of each agent has to conform to the information graph constraint, the conventional game strategy design approaches based upon Riccati equation(s) are not applicable because all the agents are required to have the information of the entire system. Accordingly, the game strategy design under distributed information is commonly known to be challenging. Toward this end, we propose novel open-loop and feedback game strategy design approaches for Nash equilibrium and noninferior solutions with a focus on linear quadratic differential games. For the open-loop design, approximate Nash/noninferior game strategies are proposed by integrating distributed state estimation into the open-loop global-information Nash/noninferior strategies such that, without global information, the distributed game strategies can be made arbitrarily close to and asymptotically converge over time to the global-information strategies. For the feedback design, we propose the best achievable performance indices based approach under which the distributed strategies form a Nash equilibrium or noninferior solution with respect to a set of performance indices that are the closest to the original indices. This approach overcomes two issues in the classical optimal output feedback approach: the simultaneous optimization and initial state dependence. The proposed open-loop and feedback design approaches are applied to an unmanned aerial vehicle formation control problem and a multi-pursuer single-evader differential game problem, respectively. Simulation results of several scenarios are presented for illustration.

To my parents and my wife, for their love and support

ACKNOWLEDGMENTS

I am grateful to Professor Zhihua Qu and Professor Marwan Simaan for giving me useful advice regarding the research in my dissertation. This dissertation would not have been possible without their support. I am also thankful to my committee members Professor Jiongmin Yong, Professor Michael Haralambous, and Professor Tuhin Das for their help.

I would like to thank Dr. Chaoyong Li, Dr. Jianan Wang, and all the other colleagues in Systems and Control Laboratory for the time they offered in valuable discussions.

Finally, I would like to thank my parents and my wife for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In this chapter, the basic background knowledge of differential game theory, multi-agent systems, and distributed information structures is covered, the motivation of this research is raised, and the scope of this dissertation is defined.

1.1   Background

1.1.1   Differential Game Theory

The game theory was originated from economics [1] in 1940s and has been widely applied in many areas such as control systems engineering, military and aerospace engineering, power systems, communication network, biomedical science, etc. It has been extensively studied and explored by many researchers in the past decades and becomes a quite mature area nowadays. The game theory basically deals with situations where two or more players are involved and making decisions to pursue their own objectives which could be their profits, performance, or utility functions in applications. Each player makes its own decision to achieve certain outcome of its objective function. This decision is called the player's strategy. A set that contains all the possible strategies from which the player can choose in the game is called the player's admissible strategy set.

During the development of the game theory, different types of games have emerged. Some typical classifications of the games are introduced as follows:

1. In terms of the time dependence of the strategy, there are static games and dynamic games.

    In a static game, every player only makes a one-shot strategy. Moreover, for a two-player static game, if every player's admissible strategy set contains a finite number of strategies,

the game is also called a matrix game. That is because the outcomes corresponding to all the possible combinations of the two players' admissible strategies can be put into a matrix form.

In a dynamic game, every player makes a strategy profile or trajectory as a function of (continuous or discrete) time for the entire game process. If the players' state dynamics are governed by differential equations, then the game is also called a differential game and the objective functions that the players try to maximize or minimize are usually called the performance indices.

2. In terms of the players' willingness to collaborate, there are noncooperative games and cooperative games.

In a noncooperative game, every player is assumed to focus on pursuing its own objective only and not to collaborate with others. A typical solution to the noncooperative game is the well-known Nash equilibrium [2]. The Nash equilibrium can be interpreted as a state where no player has the intend to unilaterally deviate from its strategy and if it does so, then a loss will occur in its objective function. Therefore, the Nash equilibrium can be regarded as a "safe" solution to prevent any player from cheating and is preferred in many applications of games with noncooperative players. The Nash equilibrium can be obtained by utilizing the standard static optimization technique for the static game and by utilizing the optimal control theory [3] for the dynamic game. Note that the existence or uniqueness of the Nash equilibrium for a noncooperative game is not always guaranteed.

In a cooperative game, although individual players have their own objective, they are assumed to collaborate with each other to jointly improve their objective functions. Therefore, a cooperative game are also regarded as a multi-objective optimization problem. A typical solution to the cooperative game is called the noninferior solution or Pareto optimality [4]. The noninferior solution can be interpreted as a state where it is impossible to improve any

player's objective function without loss in at least one player's objective function. To obtain the noninferior solution, only a single optimization problem needs be solved with the objective function being a convex combination of all the players' objective functions. For all the different choices of convex parameters, a noninferior set or Pareto frontier can be generated. Note that unlike the Nash equilibrium, the player has the intend to deviate from the noninferior solution unilaterally. As such, the noninferior solution is preferred in the situation where all the players in the game are restricted to stick with the noninferior strategy or within the same team.

3. In terms of the strategy making sequence, there are Nash games and Stackelberg games for noncooperative games. In a Nash game, all the players make decisions simultaneously and the Nash equilibrium can be obtained as we have already introduced. While in a Stackelberg game, there exists a sequence of decision making, that is, some of the players will make decisions first as the leaders and announce their strategies to the rest of the players, and the rest of the players as followers will then make decisions to react to the leaders' announced strategies. In a Stackelberg game, it is assumed that all the followers are rational and will react to the leaders' strategies in an optimal way. Knowing that the following will react optimally, the leaders will naturally announce strategies that will optimize their objective functions. The players' strategies are said to be Stackelberg strategies. Clearly, the leaders in a Stackelberg game has the advantage over the followers under the condition that they have to know the objectives of all the followers. A typical example of the Stackelberg games is the oligopoly market where there are several big dominated companies as the leaders and the other small companies as the followers. The big companies usually have information advantage over the small companies and their policies have great impact on these small companies.

4. In terms of the relationship among the players' objective functions, there are zero-sum games

and nonzero-sum games. In a zero-sum game, the sum of all the players' objective functions is equal to zero. Therefore, the total gain in some players' objective functions is equal to the total loss in the other players' objective functions. Therefore, the players have conflict objectives. In a zero-sum game, the Nash equilibrium is also known as a saddle-point solution. One typical example of the zero-sum games is the one-pursuer one-evader game where the pursuer tries to minimize the distance between the evader and itself at the terminal time while the evader tries to maximize such a distance.

In a nonzero-sum game, the sum of all the players' objective functions is not equal to zero.

In this research, since we focus on the differential games, in what follows, the historical development of the differential games is briefly introduced and the associated literatures are reviewed.

It has been commonly regarded that the introductory work on the differential game theory was done by Isaacs in 1950s with major applications in military situations and pursuit-evasion games. His book [5] focuses on the zero-sum differential games and the well-known Hamilton-Jacobi-Isaacs partial differential equation for deriving the feedback Nash Strategy was proposed. From Isaacs' pioneering work on, a plenty of results on differential games have been coming out consistently. Necessary conditions for a certain type of differential games to have a saddle point solution were derived in [6], where the calculus of variation technique was utilized first time in the differential games. One of the most important works on nonzero-sum games is [7]. This paper focuses on both static and differential nonzero-sum games. Three types of solutions were discussed, that is, the Nash equilibrium, minmax solution, and noninferior solution (later known as the solution to the cooperative game). The Nash equilibrium of the nonzero-sum linear quadratic game was obtained by solving the the coupled differential matrix Riccati equations. In [8], an important property of the linear quadratic differential games was discovered that the limiting solution of the coupled differential Riccati equations does not necessarily become the solution to the game over the infinity

time. In [9, 10, 11], the uniqueness and existence of the Nash equilibrium for linear quadratic games were studied. In parallel with the research on differential Nash games, there exist many research works on differential Stackelberg games. Static and differential Stackelberg game was investigated in [12, 13]. In [12], the Stackelberg solution was derived for linear quadratic differential games. In [13], the important property of the Stackelberg solution, the inconsistency, was discussed and hence it was shown that the well-known Bellman's optimality principle does not hold for the Stackelberg games. The existence and uniqueness of the Stackelberg solution were further studied in [14, 15, 16, 17, 18]. After 1980, the research works started focusing on the potential applications of the differential game theory to all kinds of real life applications. In [19], the idea of differential game theory was successfully applied to the H_∞ robust optimal control design problem where the designer is regarded as one player and the noise is regarded as the other player. In [20], the output consensus problem was formulated and solved under the cooperative differential game framework. In [21], a problem where a group of agents as defenders are trying to protect an asset from being destroyed by an intruder was considered and solved as a linear quadratic differential game. In [22], the online solution for the differential games was considered using the reinforcement learning. In [23], the interaction between the microgrid and main grid in the future smart grid was formulated as a discrete time Stackelberg game and the optimal generation dispatch are obtained. Among a variety of interesting applications of the differential game theory, the pursuit-evasion game has been widely studied for decades. A pursuit-evasion game basically models the process where several pursuers try to chase several evaders for a certain period of time, while the evaders try to escape at the same time. Solving a pursuit-evasion game essentially involves developing strategies for the pursuers and evaders such that their prescribed performance indices are optimized. After the pioneering work [5], the saddle point solutions for a type of zero-sum single-pursuer single-evader games were considered in [24]. Nonzero-sum pursuit-evasion games were introduced and investigated as an example of the Nash equilibrium strategies in [7] and as an example of leader-follower Stackelberg strategies in [12]. In [25], a two-pursuer one-evader game was considered. In [26],

a two-evader one-pursuer cooperative defending game was considered. In [27], the structured strategies on improving the cooperative pursuit was discussed. In [28, 29, 30], pursuit-evasion games with formation control that makes pursuers spread around the evader were studied. In [31], a homicidal chauffeur game with collaborative pursuers was discussed. In [32, 33], a derivative based strategy design approach was proposed for multi-player pursuit-evasion games. In [34], pursuit-evasion games integrating communication theory to deal with the spatial jamming problem was discussed. In [35, 36], multi-player pursuit-evasion game with evaders having higher speed than the pursuers was considered. The conventional multi-player pursuit-evasion games assume that either the pursuers or the evaders are able to have global information of the overall system, that is, every pursuer is able to observe all the other pursuers and evaders, and every evader is able to observe all the other pursuers and evaders. However, in many applications of the pursuit-evasion game, the players (either the pursuers or the evaders) might only be able to have limited information of the overall system. For instance, due to the sensing range capability or the obstacles in the environment, each player might have a limited capability to observe a subset of the players in the game. This type of multi-player pursuit-evasion games with incomplete information were investigated in [37, 38, 39, 40, 41]. A short survey [42] is recommended as a dedicated report on the pursuit-evasion games.

1.1.2 Multi-Agent Systems

As the modern system becomes more complex and large-scaled, a single system usually consist of several subsystems (or agents). This type of systems are called multi-agent systems. The control objective of such a system is to coordinate the subsystems to complete a certain task while at the same time maintaining the stability of the overall system. In most of the multi-agent systems, to achieve the coordination among the agents, there usually exists local communication or sensor networks such that the agents are able to exchange certain information with each others through

the networks. In recent years, there is a surge of research works on such systems with the potential applications in multi-vehicle coordination, signal synchronization, distributed generator control in the future smart grid, etc. In what follows, we will briefly review the research subjects and existing results on the multi-agent systems.

A large part of the research works on multi-agent systems are dedicated to the consensus problem [43, 44, 45]. This problem is essentially designing control inputs for the agents such that their outputs under these controls become identical as time goes to infinite. The control design to achieve a consensus in the multi-agent systems is also known as the cooperative control design [46] which is generally a control law that utilizes the information available to individual agents only. The consensus problem can be better illustrated by the following typical applications. A rendezvous problems is a consensus problem where a group of agents (e.g. mobile vehicles) needs be controlled to arrive at a common location in a physical environment. Note that the final rendezvous location achieved by the agents does not need to be predetermined under the typical consensus algorithm or cooperative control law. A flocking problem is a consensus problem where a group of agents (e.g. mobile vehicles) needs be controlled to achieve a common constant velocity. Note that similar to the rendezvous problem, the final common velocity achieved by the agents does not need to be predetermined under the typical consensus algorithm or cooperative control law. A formation control problem [47, 48, 49] is a consensus problem where a group of agents needs be controlled to form a prescribed formation. The formation control problem can be regarded as a combination of a rendezvous problem and flocking problem because the formation control problem is a rendezvous problem where the prescribed formation distance between any two agents is zero and the formation control problem is also a flocking problem where agents' velocities must be identical to preserve the formation. A synchronization problem [50, 51, 52, 53] is a consensus problem where the agents' outputs need be controlled to track a prescribed reference trajectory. Unlike the previous rendezvous, flocking, and formation control problems where the consensus value is

identical and constant as times goes to infinity, the reference trajectory as the consensus value of the synchronization problem is usually time-varying. Synchronization problems has great potential applications in power generation industry in terms of synchronization of the voltage, phase, or frequency among a large number of distributed generators in the future smart grid. As we can clearly see from the above applications, the consensus is a "stable" status of a multi-agent system, however, is different from the conventional concept of stability where all the states or outputs of the agents vanish as the time goes to infinite. Therefore, to distinguish with the conventional concept, the stability for the consensus of multi-agent systems is called the cooperative stability, which is a status where an agreement on all the agents' states of interest is achieved.

Another research area is dedicated to the optimal control design in the multi-agent systems. The optimal control design for a multi-agent system is essentially finding the control input compatible with the communication or sensor network for each and every agent such that a given performance index is minimized. Due to the communication or sensor network constraint, the classical optimal control design approach [3] is generally not applicable. For instance, since the linear feedback control for each agent has to be structured, the well-known Riccati equation approach for the linear quadratic optimal control is not applicable because the feedback gain matrix obtained by this approach is a full matrix in general. As such, a different optimal control design approach must be proposed for multi-agent systems. In what follows, we briefly introduce major research directions toward solving this problem (with the focus on linear systems). There are approaches based on optimal output feedback control design. Since the linear feedback control in a multi-agent system can be treated as multiple control inputs with different output feedback channels, the problem can be solved under the framework of optimal output feedback control design. The pioneering works on the optimal output feedback control design are [54, 55] where the basic idea is to parameterize the gain matrix and optimize it directly with respect to the given performance index. In these papers, an gradient based iterative algorithm for computing the optimal feedback

matrix was proposed for the finite and infinite time horizon. The computational complexity of this algorithm was later shown to be NP-hard in [56]. A comprehensive survey on the optimal output feedback control was included in [57]. Applying the optimal output feedback control design to the multi-agent systems was discussed in [58] in terms of optimal decentralized control design and a numerical algorithm similar to the one in [54] was proposed. There are approaches based on the transformation technique. In [59], the optimal decentralized control for a string of vehicles was derived using the spatial transformation technique, where the dynamics and the information exchange pattern were assumed to be identical for every vehicle. In [60, 61], the transformation technique was further explored and the property of identical agents' dynamics and information exchange pattern was defined as spatially invariance. There are approaches based on the convex optimization technique. Since the optimal control design for multi-agent systems is generally a non-convex problem, the conventional convex optimization tools cannot be applied. In [62, 63], this non-convex problem was recast as a convex problem under rather strict conditions. Other approaches include the graph approach [64] and linear quadratic approach for identical systems [65]. The optimal control problem for multi-agent system has been investigated for a long time and commonly regarded as a very hard problem.

There are research works that utilize the differential game theory to solve the multi-agent control problem. In [66], the formation control problem was formulated as a noncooperative differential game and the receding horizon Nash equilibrium was solved. In [67], the consensus problem was formulated as a cooperative differential game and the Nash bargain solution among the Pareto-efficient solutions was found using linear matrix inequality (LMI) approach. In [68], a zero-sum game was formulated between the sensor network and an intelligent moving target, and a robust target position estimator was obtained.

There are many other research areas on multi-agent systems, including robust control, time-delay control, network optimization, etc, which cannot be fully covered in this dissertation. A compre-

hensive review on the multi-agent systems can be found in [69].

1.1.3   Distributed Information Structures

An important factor of a differential game is each player's information structure. In a differential game, each player's information structure is extremely important because its strategy profile resulted from the different information structures can be completely different. There are two typical information structures. One is the open-loop information structure and the other is the feedback information structure. Consequently, we call the open-loop strategy for the player's strategy under open-loop information structure and feedback strategy for the player's strategy under feedback information structure.

In the conventional game, the information available to the player is assumed to be "global" where every player is assumed to have the information of all the other players in the game. Therefore, in this case, the player is under open-loop information structure if only the global information of the system at the initial time is available to it along the game process as shown in Figure 1.1.

Figure 1.1: Global open-loop information structure

The player is under feedback information structure if the global information of the system is avail-

able to it at every instant of time along the game process as shown in Figure 1.2.
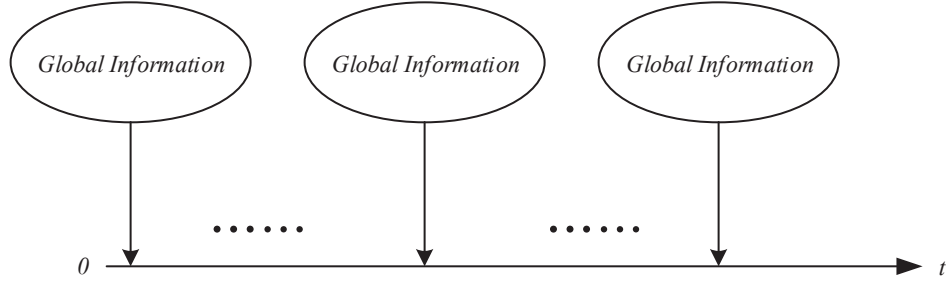
Figure 1.2: Global feedback information structure

If the game takes place in a large-scale system such as a multi-agent system, each agent is only able to have the information of a subset of (possibly neighboring) agents through the local communication or sensor networks. In this situation, since the global information is not available any more, the players are said to have distributed information. Therefore, we can extend the concepts of open-loop and feedback information structures under distributed information. The distributed open-loop information structure is shown in Figure 1.3 where the blue subset inside the circle (the global information) stands for the distributed information available to the player.

Figure 1.3: Distributed open-loop information structure

11

The distributed feedback information structure is shown in Figure 1.4 where the blue subset inside the circle (the global information) stands for the distributed information available to the player.
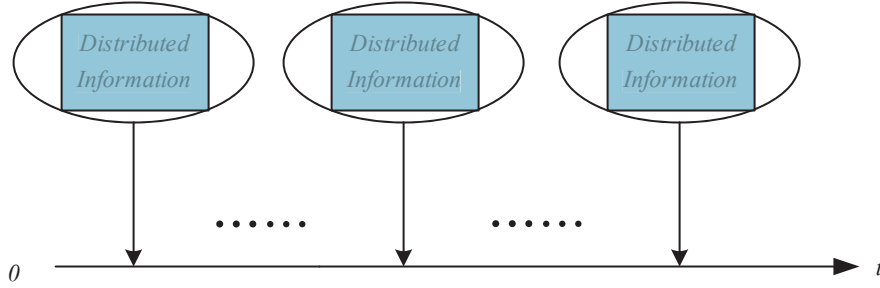
Figure 1.4: Distributed feedback information structure

Moreover, the strategy under distributed open-loop information structure is called the distributed open-loop strategy and the strategy under distributed feedback information is called the distributed feedback strategy.

1.2   Motivation and Scope

The motivation of this dissertation lies in the following two aspects.

1. The first aspect is that most of the existing results on differential games assume that every player has global information and the games under distributed information have not been well studied, which is in fact quite common and important in many applications involving large-scale systems, such as multi-agent systems.

2. The second aspect is that in conventional multi-agent control design problem, the agents are assumed to pursue a goal of optimizing a common performance index. However, it is of

practical interest to consider a situation where individual agents try to optimize their own objective functions.

The above two aspects actually lead to considering the differential game problem for multi-agent systems under distributed information. So far, there are very few research works in this area. Therefore, it is in a great demand to propose a strategy design approach such that each agent only utilizes the information available to it. In this dissertation, we will focus on the linear quadratic differential games in the multi-agent system where the dynamics of each agent is governed by a linear differential equation and the performance index of each agent is in the quadratic form. We will consider the design approaches of Nash equilibrium and noninferior solutions under both distributed open-loop and feedback information structures. The remainder of this dissertation is organized as follows:

In Chapter 2, the linear quadratic differential games in the multi-agent system under distributed information is formulated.

In Chapter 3, the open-loop strategy design based upon Riccati equation is introduced first. A distributed strategy design approach is then proposed by integrating a novel distributed state estimation law.

In Chapter 4, the feedback strategy design based upon Riccati equation and the classical design based upon optimal output feedback control are introduced first. A distributed strategy design approach is then proposed based on a novel concept of the best achievable performance indices.

In Chapter 5, the proposed approaches are applied to an unmanned aerial vehicle formation control problem and a multi-pursuer single-evader differential game problem with limited observations.

In Chapter 6, the dissertation is concluded with the summary of the results obtained in this research and the future research directions.

CHAPTER 2: PROBLEM FORMULATION

In this chapter, the linear quadratic differential game in an N-agent system is formulated and formal definitions of related concepts are given.

2.1 System Dynamics

There are N agents who have decoupled linear dynamics and given by

$$\dot{x}_i(t) = A_i(t)x_i(t) + \hat{B}_i(t)u_i(t) \tag{2.1a}$$

$$y_i(t) = C_i(t)x_i(t), \tag{2.1b}$$

for $i = 1, \cdots, N$, where $x_i \in \mathbb{R}^{n_i}$ is the state vector, $u_i \in \mathbb{R}^{m_i}$ is the control input, $y_i \in \mathbb{R}^r$ is the output vector. Matrices $A_i(t)$, $\hat{B}_i(t)$, $C_i(t)$ are time-varying and of proper dimensions. Since each agent is an independent entity in real life applications, we assume that $\{A_i, B_i\}$ is a controllable pair, matrix B_i is of full column rank (meaning no redundant input), and matrix C_i is of full row rank (meaning no redundant output). Agent i's initial state is given by $x_{i0} = x_i(0)$. Note that all the agents' outputs have the same dimension, which is normally required by most of the multi-agent system applications such as formation control, synchronization, pursuit-evasion games, etc. Denoting $x = [x_1^T \; \cdots \; x_N^T]^T$ and $y = [y_1^T \; \cdots \; y_N^T]^T$, the overall system can be expressed more compactly as

$$\dot{x}(t) = A(t)x(t) + \sum_{j=1}^{N} B_j(t)u_j(t) \tag{2.2a}$$

$$y(t) = C(t)x(t), \tag{2.2b}$$

where

$$
A(t) = \begin{bmatrix} A_1(t) & & \\ & \ddots & \\ & & A_N(t) \end{bmatrix}, \quad B_i(t) = \begin{bmatrix} 0_{n_1 \times m_i} \\ \vdots \\ \hat{B}_i(t) \\ \vdots \\ 0_{n_N \times m_i} \end{bmatrix}, \quad C(t) = \begin{bmatrix} C_1(t) & & \\ & \ddots & \\ & & C_N(t) \end{bmatrix}, \quad (2.3)
$$

where $0_{n_j \times m_i}$ is the $n_j \times m_i$ zero matrix for $j = 1, \cdots, N$ and matrix $\hat{B}_i(t)$ is the ith block of matrix $B_i(t)$.

2.2 Information Structure

The agents in the system is able to exchange information with others through the communication/sensor network. This information exchange pattern or information flow among the agents is often described by a directed information graph denoted by $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ where node $v_i \in \mathcal{V}$ represents agent i for $i = 1, \cdots, N$ and edge $e_{ij} \in \mathcal{E}$ represents the directional information flow from node j to node i (it is always true that $e_{ii} \in \mathcal{E}$ since agent i can always have its own information). If the information exchange pattern is fixed over time, then the graph is fixed. If the information exchange pattern changes over time (due to possible communication failure or obstacles in the environment, etc.), then the graph is time-varying. In this dissertation, we primarily consider the fixed information exchange pattern and hence we make the following assumption.

Assumption 2.1. *The information graph in the multi-agent system is fixed.*

For example, Figure 2.1 shows a fixed information graph among four agents.

16
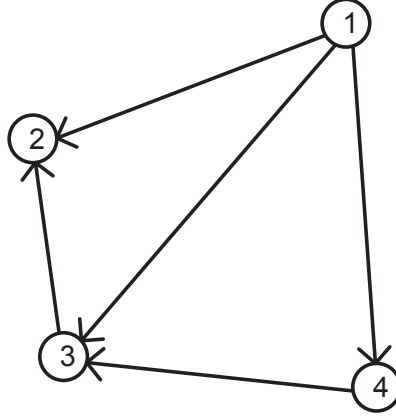
Figure 2.1: Information graph among four agents

Clearly, in this graph, $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$ and $\mathcal{E} = \{e_{11}, e_{22}, e_{33}, e_{44}, e_{21}, e_{31}, e_{41}, e_{23}, e_{34}\}$. Several important concepts related with the information graph are introduced as follows:

Definition 2.1 (Path). *In directed graph $\mathcal{G}$, a path from node v_i to node v_j is a sequence of directed edges that connect from node i to node j.*

Based on the definition of path, the globally reachable node can be defined as follows:

Definition 2.2 (Globally Reachable Node). *In directed graph $\mathcal{G}$, node v_i is globally reachable if there exist paths from node v_i to node v_j for all $j = 1, \cdots, N, j \neq i$.*

In some literature, the globally reachable node is also regarded as the root node of the spanning tree of the graph. Furthermore, based on the definition of globally reachable node, the graph connectivity is defined as follows:

Definition 2.3 (Connected Graph). *Directed graph $\mathcal{G}$ is connected if it contains at least one globally reachable node.*

One important matrix associated in the graph theory is the Laplacian matrix [70]. This matrix is denoted as $\mathcal{L} = [\mathcal{L}_{ij}]$, where

$$
\mathcal{L}_{ij} = \begin{cases} -l_{ij} & \text{if } e_{ij} \in \mathcal{E} \text{ and } i \neq j \\ \sum_{j=1,j\neq i}^{N} l_{ik} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad , \tag{2.4}
$$

where l_{ij} is a positive scalar. For example, the Laplacian matrix associated with the graph in Figure 2.1 is given by

$$
\mathcal{L} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -l_{21} & l_{21}+l_{23} & -l_{23} & 0 \\ -l_{31} & 0 & l_{31}+l_{34} & -l_{34} \\ -l_{41} & 0 & 0 & l_{41} \end{bmatrix}.
$$

The Laplacian matrix has numerous well-known properties, two of which are presented as follows:

Proposition 2.1. *The Laplacian matrix has the following properties:*

1. *If a directed graph is connected, then the null space of the associated Laplacian matrix is spanned by a single vector $\boldsymbol{1}_N$, where $\boldsymbol{1}_N$ is an $N \times 1$ vector with all the entries equal to 1.*

2. *All the eigenvalues of the Laplacian matrix associated with any directed graph have non-negative real parts.*

The above concepts of the graph theory in fact describe the distributed information among the agents. For the differential game in the N-agent system, the distributed open-loop and feedback information structure can be formally defined as follows:

- agent i is under the **open-loop** information structure if its strategy at time t can only depends

18

on the initial output of agent j, $y_j(0)$, for all j such that $e_{ij} \in \mathcal{E}$ on the directed information graph $\mathcal{G}$.

- agent i is under the **feedback** information structure if its strategy at time t can only depend on the output of agent j, $y_j(t)$, at the current time t for all j such that $e_{ij} \in \mathcal{E}$ on the directed information graph $\mathcal{G}$.

2.3 Performance Indices

In the multi-agent system, each agent tries to minimize its own performance index. In this dissertation, we consider the following performance indices for the agents:

$$J_i = \frac{1}{2}\|y(t_f)\|_{F_i}^2 + \frac{1}{2}\int_0^{t_f}[\|y(t)\|_{Q_i(t)}^2 + \|u_i(t)\|_{R_i(t)}^2]\mathrm{d}t \quad \forall i = 1, \cdots, N. \tag{2.5}$$

where where t_f is the terminal time of the game, $\|u_i(t)\|_{R_i(t)}^2 = u_i^T(t)R_i(t)u_i(t)$, and matrix $R_i(t)$ is positive definite to ensure the convexity. Matrices F_i and Q_i are usually positive semi-definite, however, some exception exists (e.g. pursuit-evasion games). The above performance indices can be utilized to characterize a plenty of applications in multi-agent systems. Several typical applications of the multi-agent system characterized by performance index (2.5) with the different choices of matrices F_i and Q_i are presented and explained as follows.

1. **Output Regulation**. In this application, matrices F_i, Q_i, and R_i are positive definite for all $i = 1, \cdots, N$. A typical choice of the coefficients for the output regulation that takes the information graph into account is

$$\|y(t_f)\|_{F_i}^2 = \sum_{e_{ij} \in \mathcal{E}} f_{ij}\|y_j(t_f)\|_2^2 \quad \text{and} \quad \|y(t)\|_{Q_i(t)}^2 = \sum_{e_{ij} \in \mathcal{E}} q_{ij}(t)\|y_j(t)\|_2^2.$$

where $\| \cdot \|_2$ stands for the Euclidean norm and $f_{ij}, q_{ij}(t)$ are positive scalars for all $j = 1, \cdots, N$ and all $t \in [0, t_f]$. In this case, performance index (2.5) essentially means that agent i tries to drive the entire output vector to zero while at the same time minimizing its control effort over the entire process.

2. **Output Consensus**. In this application, matrices F_i and Q_i are positive semi-definite for all $i = 1, \cdots, N$ and matrix R_i is positive definite for all $i = 1, \cdots, N$. A typical choice of the coefficients for the output consensus that takes the information graph into account is

$$\|y(t_f)\|_{F_i}^2 = \sum_{e_{ij} \in \mathcal{E}} f_{ij} \|y_i(t_f) - y_j(t_f)\|_2^2 \quad \text{and} \quad \|y(t)\|_{Q_i(t)}^2 = \sum_{e_{ij} \in \mathcal{E}} q_{ij}(t) \|y_i(t) - y_j(t)\|_2^2.$$

In this case, performance index (2.5) essentially means that agent i tries to drive all the output vectors (or part of the output vectors) to a common value while at the same time minimizing its control effort over the entire process.

3. **Multi-Pursuit Single-Evasion Game**. Suppose that the output y_i stands for agent i's position and agents 2 to N (pursuers) try to chase agent 1 (evader) who tries to evade the pursuers. In this situation, the performance indices of the agents are given by (2.5) where matrices F_1 and Q_1 are negative definite, matrix R_1 is positive definite, matrices F_i and Q_i are positive definite for all $i = 2, \cdots, N$, and matrix R_i is positive definite for all $i = 2, \cdots, N$. A typical choice of the coefficients is

$$\|y(t_f)\|_{F_i}^2 = f_i \|y_i(t_f) - y_1(t_f)\|_2^2, \quad \|y(t_f)\|_{F_1}^2 = -\sum_{j=1}^N f_{1j} \|y_1(t_f) - y_j(t_f)\|_2^2,$$

$$\|y(t)\|_{Q_i(t)}^2 = q_i(t) \|y_i(t) - y_1(t)\|_2^2, \quad \|y(t)\|_{Q_1(t)}^2 = -\sum_{j=1}^N q_{1j}(t) \|y_1(t) - y_j(t)\|_2^2,$$

where f_i, q_i, f_{1j}, and q_{1j} are positive scalars for all $j = 2, \cdots, N$. In this case, performance index J_1 essentially means that the evader tries to maximize its distances to the pursuers

while at the same time minimizing its control effort over the entire process and on the other hand, performance index J_i for all $i = 2, \cdots, N$ essentially means that the pursuer i tries to minimize its distances to the evader while at the same time minimizing its control effort over the entire process.

2.4   Game Solutions

Given the system equation in (2.2) and performance indices in (2.5), a differential game problem is formulated. In this dissertation, we consider both the noncooperative game and the cooperative game for the multi-agent system and the different types of solutions to these games are introduced as follows. If the differential game for the N-agent system is noncooperative, the Nash equilibrium [2] is defined as follows:

Definition 2.4 (Nash Equilibrium)**.** *For the differential game in the N-agent system defined by system dynamics in (2.2) and performance indices in (2.5), the strategies $u_1^*, \cdots, u_N^*$ form a Nash equilibrium if the inequalities*

$$J_i(u_1^*, \cdots, u_N^*) \leq J_i(u_1^*, \cdots, u_{i-1}^*, u_i, u_{i+1}^*, u_N^*) \quad \forall u_i \in U_i \quad \forall i = 1, \cdots, N \qquad (2.6)$$

hold, where U_i is agent i's admissible strategy set.

Moreover, the ε-Nash equilibrium for the noncooperative differential game for the N-agent system is defined as follows:

Definition 2.5 (ε-Nash equilibrium)**.** *Given the differential game in an N-agent system defined by system dynamics in (3.11) and performance indices in (5.9) and a real non-negative parameter ε, the agents' strategies are said to form an ε-Nash equilibrium if it is not possible for any agent*

21

to reduce more than ε in its performance index value by unilaterally deviating from its strategy.
Formally, strategies $u_1, \cdots, u_N$ form an ε-Nash Equilibrium if the inequalities

$$J_i(u_1^*, \cdots, u_N^*) \leq J_i(u_1^*, \cdots, u_{i-1}^*, u_i, u_{i+1}^*, \cdots, u_N^*) + \varepsilon \quad \forall u_i \in U_i, i = 1, \cdots, N. \qquad (2.7)$$

hold, where U_i is agent i's admissible strategy set.

It is clear that every Nash Equilibrium is equivalent to an ε-Nash equilibrium where $\varepsilon = 0$. If the differential game for the N-agent system is cooperative, the noninferior solution (also known as the Pareto optimality solution) is defined as follows:

Definition 2.6 (Noninferior Solution)**.** *For the differential game in an N-agent system defined by system dynamics in (2.2) and performance indices in (2.5), the strategies $u_1^*, \cdots, u_N^*$ form a noninferior solution if there exists $i \in \{1, \cdots, N\}$ such that the inequalities*

$$J_i(u_1, \cdots, u_N) \leq J_i(u_1^*, \cdots, u_N^*) \quad \forall u_1 \in U_1, \cdots, u_N \in U_N \quad \forall i = 1, \cdots, N \qquad (2.8)$$

*do **not** hold with at least one strict inequality, where U_i is agent i's admissible strategy set.*

The noninferior solution can be interpreted as a solution in which any changes made do not help improve every agent's performance index value.

CHAPTER 3: OPEN-LOOP GAME STRATEGIES

In this chapter, we consider the open-loop game strategy design for both the Nash equilibrium and noninferior solution in the formulated N-agent system. We will first introduce the existing Riccati equation approach and then present the proposed approach based on a distributed state estimation algorithm.

3.1   Riccati Equation Approach

To derive the open-loop game strategy, we utilize the well-known Pontryagin's minimum principle [71] which yields the necessary optimality conditions for both the Nash equilibrium strategy and noninferior solution strategy. These two types of strategies are presented as follows:

Open-loop Nash Equilibrium Strategy: The open-loop Nash equilibrium under the linear quadratic framework is well-known [7] and the result is presented as the following theorem without proof.

Theorem 3.1. *For the differential game in an N-agent system defined by system dynamics in (2.2) and performance indices in (2.5), the strategies*

$$u_i(t) = -R_i^{-1}(t)B_i^T(t)P_i(t)\phi(t,0)x(0) \quad \forall i = 1, \cdots, N \tag{3.1}$$

form an open-loop Nash equilibrium, where matrix $\phi(t,0)$ is the closed-loop state transition matrix defined by $\phi(t,0) = e^{\bar{A}t}$ and $\bar{A} = A - \sum_{j=1}^{N} B_j R_j^{-1} B_j^T P_j$, and matrix $P_i(t)$ is the solution to the following coupled differential Riccati equations

$$\dot{P}_i + P_i A + A^T P_i - P_i \sum_{j=1}^{N} B_j R_j^{-1} B_j^T P_j + C_i^T Q_i C_i = 0 \quad \forall i = 1, \cdots, N \tag{3.2}$$

with the boundary condition $P_i(t_f) = C^T(t_f)F_iC(t_f)$.

Open-loop Noninferior Solution Strategy: The noninferior solution strategy is essentially de-rived by minimizing the following convex combination of all the agents' performance indices

$$J = \sum_{j=1}^{N} \alpha_j J_j \tag{3.3}$$

where $0 \leq \alpha_j \leq 1$ and $\sum_{j=1}^{N} \alpha_j = 1$. This minimization is essentially a optimal control prob-lem with parameters $\alpha_1, \cdots, \alpha_N$. The open-loop noninferior solution under the linear quadratic framework was obtained in [7] and the result is presented as the following theorem without proof.

Theorem 3.2. *For the differential game in an N-agent system environment defined by system dy-namics in (2.2) and performance indices in (2.5), the strategies*

$$u_i(t) = -\frac{1}{\alpha_i}R_i^{-1}(t)B_i^T(t)P(t)\phi(t,0)x(0) \quad \forall i = 1, \cdots, N \tag{3.4}$$

form an open-loop noninferior solution, where matrix $\phi(t,0)$ is the closed-loop state transition matrix defined by $\phi(t,0) = e^{\bar{A}t}$ and $\bar{A} = A - \sum_{j=1}^{N} \frac{1}{\alpha_j}B_j R_j^{-1}B_j^T P$, and matrix $P(t)$ is the solution to the following differential Riccati equation

$$\dot{P} + PA + A^T P - \sum_{j=1}^{N} \frac{1}{\alpha_j}PB_j R_j^{-1}B_j P + \sum_{j=1}^{N} \alpha_j C_j^T Q_j C_j = 0 \tag{3.5}$$

with the boundary condition $P(t_f) = \sum_{j=1}^{N} \alpha_j C_j^T(t_f)F_j C_j(t_f)$.

Note that the above approaches requires to solve for the matrix $P_i(t)$ from the coupled differential Riccati equations in (3.1) backward in time or matrix $P(t)$ from the differential Riccati equation in (3.5) backward in time. After solving the differential equation (usually with the aid of computer),

24

the matrix solution will generally become a full matrix with all the entries being nonzero. There-fore, by looking at the expression in (3.1) or (3.4), individual agents in the game needs to have the complete knowledge of the initial state information, $x(0)$, in order to implement their open-loop Nash equilibrium strategies or noninferior solution strategies. In the conventional game problem, this requirement has no problem because it is always assume that all the required information is available for each and every player. However, if the differential game takes place in a multi-agent system under distributed information, then agent i can acquire the information of the agent j only if $e_{ij} \in \mathcal{E}$ and hence is not able to implement the strategy (3.1) or (3.4) derived using the Riccati equation(s) approach.

3.2 Distributed Game Strategy Design

Realizing that the open-loop Nash equilibrium and noninferior solution expressed in (3.1) and (3.4) are not implementable in the multi-agent system under distributed information, a new approach for the open-loop Nash equilibrium and noninferior solution design must be proposed for the agents such that they can carry it out to accommodate the distributed information. To achieve this, first of all, the performance needs be well structured according to the information graph among the agents. We define a block diagonal matrix $D_i \in \mathbb{R}^{Nr \times Nr}$ as follows:

$$D_i = \sum_{e_{ij} \in \mathcal{E}} (d_j d_j^T) \otimes I_r \quad \forall i = 1, \cdots, N, \tag{3.6}$$

where $\otimes$ is the Kronecker product and I_r is the $r \times r$ identity matrix. The the jth diagonal block is equal to I_r if $e_{ij} \in \mathcal{E}$ and 0 if $e_{ij} \notin \mathcal{E}$. For instance, the matrices D_1, D_2, D_3, D_4 for the information

graph shown in Figure 2.1 are

$$
D_1 = \begin{bmatrix} I_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} I_r & 0 & 0 & 0 \\ 0 & I_r & 0 & 0 \\ 0 & 0 & I_r & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad D_3 = \begin{bmatrix} I_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_r & 0 \\ 0 & 0 & 0 & I_r \end{bmatrix}, \quad D_4 = \begin{bmatrix} I_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_r \end{bmatrix},
$$

respectively. The product of $(D_i y)$ is a vector with the jth block entry to y_j if $e_{ij} \in \mathcal{E}$ and 0 if $e_{ij} \notin \mathcal{E}$. Therefore, we consider the structured performance index (2.5) with matrices

$$
F_i = D_i^T \tilde{F}_i D_i \tag{3.7}
$$

and $Q_i = 0$, which can be expressed as

$$
J_i = \frac{1}{2} \|y(t_f)\|_{D_i^T \tilde{F}_i D_i}^2 + \frac{1}{2} \int_0^{t_f} \|u_i(t)\|_{R_i(t)}^2 \mathrm{d}t \quad \forall i = 1, \cdots, N. \tag{3.8}
$$

where matrix $\tilde{F}_i$ can be selected appropriately to achieve desired control objective of agent i as discussed in Section 2.3. Basically, performance index (3.8) means that agent i tries to minimize a cost term that only involves the outputs of agent j for all $e_{ij} \in \mathcal{E}$ at the terminal time while at the same time minimizing its control effort over the entire game process. For the differential game defined by system (2.2) and performance indices (3.8), it will turn out that the open-loop Nash equilibrium and noninferior solution have explicit expressions and can be utilized to construct a distributed strategy synthesis algorithm under certain condition. First of all, we define a new state vector as

$$
z_i = C_i(t)\phi_i(t_f, t)x_i(t) \tag{3.9}
$$

where $\phi_i(t_f, t) = e^{A_i(t_f-t)}$. Differentiating the above equation with respect to t and recalling system dynamics in (2.1) yields:

$$\dot{z}_i = C_i \dot{\phi}_i x_i + C_i \phi_i \dot{x}_i = B_i'(t)u_i, \tag{3.10}$$

where $B_i'(t) = C_i \phi_i(t_f, t)\hat{B}_i$. Denoting $z = [z_1^T \ \cdots \ z_N^T]^T$, the system dynamics can be expressed more compactly as follows:

$$\dot{z} = \begin{bmatrix} B_1' u_1 \\ \vdots \\ B_N' u_N \end{bmatrix} = \sum_{j=1}^{N} \tilde{B}_j u_j \tag{3.11}$$

where $\tilde{B}_j(t) = (d_j \otimes I_r)B_j'(t)$ and d_j is an $N \times 1$ vector with the jth entry equal to 1 and the other entries equal to 0. Since it is clear that $z(t_f) = y(t_f)$, performance indices in (3.8) can be also expressed as

$$J_i = \frac{1}{2}\|z(t_f)\|_{F_i}^2 + \frac{1}{2}\int_0^{t_f} \|u_i(t)\|_{R_i(t)}^2 dt \quad \forall i = 1, \cdots, N. \tag{3.12}$$

where matrix F_i is as defined in (3.7).

3.2.1 Nash Strategy Design

The open-loop Nash equilibrium can be derived using Pontryagin's minimum principle and is presented as follows.

Theorem 3.3. *For the differential game in an N-agent system under distributed information de-*

fined by system dynamics in (3.11) and performance indices in (5.9), the strategies

$$u_i = - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i M^{-1} z(0) \tag{3.13}$$

for $i = 1, \cdots, N$ form an open-loop Nash equilibrium if matrix M defined by

$$M = I_{Nr} + \sum_{j=1}^{N} S_j F_j, \tag{3.14a}$$

$$S_j = \int_0^{t_f} \tilde{B}_j(t) R_j^{-1}(t) \tilde{B}_j^T(t) dt, \tag{3.14b}$$

is invertible.

Proof. We define the Hamiltonian for agent i as

$$H_i = \frac{1}{2} \|u_i\|_{R_i}^2 + \lambda_i^T \sum_{j=1}^{N} \tilde{B}_j u_j$$

where vector $\lambda_i \in \mathbb{R}^{Nr}$ is the Lagrangian multiplier. Since the second order partial derivative of H_i with respect to u_i is equal to R_i and hence is positive define, the following conditions for u_i to minimize the performance index are necessary and sufficient

$$\dot{z} = \frac{\partial H_i}{\partial \lambda_i} = \sum_{j=1}^{N} \tilde{B}_j u_j, \tag{3.15a}$$

$$\dot{\lambda}_i = -\frac{\partial H_i}{\partial z} = 0, \quad \lambda_i(t_f) = F_i z(t_f), \tag{3.15b}$$

$$\frac{\partial H_i}{\partial u_i} = R_i u_i + \tilde{B}_i^T \lambda_i = 0. \tag{3.15c}$$

Condition (3.15b) indicates that λ_i is a constant vector and hence

$$\lambda_i(t) = F_i z(t_f) \quad \forall t \in [0, t_f]$$

From (3.15c), we obtain

$$u_i = -R_i^{-1}\tilde{B}_i^T \lambda_i = -R_i^{-1}\tilde{B}_i^T F_i z(t_f) = -R_i^{-1}\tilde{B}_i^T D_i^T \tilde{F}_i D_i z(t_f). \qquad (3.16)$$

Substituting (3.16) into (3.15a) and integrating both sides from 0 to t_f yield

$$\left(I_{Nr} + \sum_{j=1}^N S_j F_j\right) z(t_f) = z(0)$$

$$M z(t_f) = z(0) \qquad (3.17)$$

where matrices M and S_j are defined in (3.14). Therefore, if the matrix M is invertible, then

$$z(t_f) = M^{-1} z(0). \qquad (3.18)$$

Substituting (3.18) into (3.16) yields the open-loop Nash strategy (3.13). $\qquad\square$

Note that M^{-1} in (3.13) is generally a full matrix. Therefore, implementing open-loop Nash strategy u_i in (3.13) will still requires every agent to have the complete knowledge of the initial state $z(0) = [z_1^T(0), \cdots, z_N^T(0)]^T$. However, it is worthwhile noting that if we define a new vector $\tilde{z} = [\tilde{z}_1^T \ \cdots \ \tilde{z}_N^T]^T$ such that $\tilde{z} = M^{-1} z(0)$, then the strategy u_i expressed in (3.13) has an interesting distributed property, that is, agent i is able to implement (3.13) as long as agent j sends its the information of $\tilde{z}_j$ to agent i for all $e_{ij} \in \mathcal{E}$ because the product $D_i M^{-1} z(0) = D_i \tilde{z} = \sum_{e_{ij} \in \mathcal{E}} d_j \otimes \tilde{z}_j$ in (3.13) only needs the information of $\tilde{z}_j$ for all $e_{ij} \in \mathcal{E}$. Therefore, in order for the agents to obtain the value of $\tilde{z}$ in a distributed manner, the basic idea is to let the agents exchange the estimates denoted by $z_1^f, \cdots, z_N^f$ through the communication network and make these estimates asymptotically converge to the actual value of $\tilde{z}_1, \cdots, \tilde{z}_N$. Toward that end, we have the following result.

Theorem 3.4. *For the differential game in an N-agent system environment defined by system dynamics in (3.11) and performance indices in (5.9), if*

1. *matrix $(-M)$ in (3.14a) is Hurwitz (all the eigenvalues have negative real parts) and*

2. *agent i updates its state estimate, $z_i^f \in \mathbb{R}^r$, for all $i = 1, \cdots, N$ according to*

$$\dot{z}_i^f = g \left[z_i(0) - z_i^f - (d_i^T \otimes I_r) S_i D_i^T \tilde{F}_i D_i z^f \right] \tag{3.19}$$

from any initial condition $z_i^f(0)$, where g is a positive scalar and $z^f = [(z_1^f)^T \ \cdots \ (z_N^f)^T]^T$,

then

$$\lim_{t \to \infty} z_i^f(t) = (d_i^T \otimes I_r) M^{-1} z(0) = \tilde{z}_i(t_f). \tag{3.20}$$

Proof. Stacking equation (3.19) from $i = 1$ to $i = N$ yields

$$\begin{aligned} \dot{z}^f &= g \left[z(0) - z^f - \sum_{j=1}^{N} S_j F_j z^f \right] \\ &= g \left[z(0) - M z^f \right]. \end{aligned} \tag{3.21}$$

where matrix M is defined in (3.14a). If matrix $(-M)$ is Hurwitz, linear system (3.21) with respect to z^f is asymptotically stable starting from any initial condition $z^f(0)$ and z^f asymptotically converges to the equilibrium of the differential equation (3.21), that is,

$$\lim_{t \to \infty} z^f(t) = M^{-1} z(0).$$

Multiplying $(d_i^T \otimes I_r)$ on both sides of the above equation yields (3.20), indicating that agent i's state estimate, $z_i^f(t)$, converges to $\tilde{z}_i$ as t goes to infinity. □

Note that to carry out the estimation law (3.19), agent i only needs to

- **retain** its private information: $z_i(0)$, S_i, F_i,

- **send** its state estimate z_i^f to agent j if $e_{ji} \in \mathcal{E}$, and

- **receive** the state estimate(s) z_j^f from agent j for all $e_{ij} \in \mathcal{E}$ because the product $(D_i z^f)$ in (3.19) is a function of z_j^f for all $e_{ij} \in \mathcal{E}$ only.

Since it is better to make the distributed state estimation algorithm (3.19) converge fast, one can increase the positive scalar g to achieve satisfactory convergence speed. Note that to successfully implement the above state estimation law (3.19), a critical condition is that all the eigenvalues of matrix M defined in (3.14a) have to have positive real parts. This condition appears to be quite stringent, however, this condition can be satisfied in many multi-agent system applications, such as the rendezvous problem and formation control problem. Note that one important feature of the proposed algorithm is that to implement it, every agent does not need to know the other agents' system dynamics, performance indices, or the overall graph connection. This fully distributed feature of the proposed approach is preferred in many real life applications.

With this distributed state estimation law, one possible way to implement the open-loop Nash strategy is to let all the agents in the system communicate for a while until a satisfactory convergent value of the state, say $\bar{z}^f$, is reached before the game starts. The agents will then implement the open-loop Nash strategy expressed in (3.13) with $D_i M^{-1} z(0)$ replaced by $D_i \bar{z}^f$. Such a design approach can be regarded as an offline computation among the agents. Although the offline approach provides accurate enough open-loop Nash strategy, it may not be applicable to the situation that requires the real-time implementation. To overcome this issue, combining the open-loop Nash strategy expressed in (3.16) along with the state estimation algorithm (3.19), an online open-loop

Nash strategy design algorithm is proposed as follows:.

$$\dot{z}_i^f = g\left[z_i(0) - z_i^f - (d_i^T \otimes I_r)S_i D_i^T \tilde{F}_i D_i z^f\right] \tag{3.22a}$$

$$u_i = -R_i^{-1}\tilde{B}_i^T D_i^T \tilde{F}_i D_i z^f \tag{3.22b}$$

for all $i = 1, \cdots, N$. Since differential equation (3.22a) is asymptotically stable, the strategy (3.22b) is actually an approximate of the actual open-loop Nash strategy in the early (transient) stage of the game and becomes sufficiently close to the actual open-loop Nash strategy thereafter. Therefore, the inequalities (2.6) in the definition of Nash equilibrium in fact does not hold under this online computing strategy and the agents can have the intend to deviate unilaterally. To quantify the agents' willingness to unilaterally deviate from the proposed strategy in (3.22), we utilize the concept of ε-Nash equilibrium in Definition 2.5. In what follows, the proposed online computing strategies in (3.22) will be shown to form an ε-Nash equilibrium and the value of ε will be derived. First of all, we present the following lemmas:

Lemma 3.1. *All the eigenvalues of matrix*

$$S_{jj} = (d_j^T \otimes I_r)S_j D_j^T \tilde{F}_j (d_j \otimes I_r) \tag{3.23}$$

has nonnegative real parts for all $j = 1, \cdots, N$ where matrix S_j is defined in (3.14b).

Proof. Substituting (3.14b) into (3.23) yields

$$S_{jj} = (d_j^T \otimes I_r)S_j(d_j \otimes I_r) = (d_j^T \otimes I_r)\int_0^{t_f} \tilde{B}_j(t)R_j^{-1}(t)\tilde{B}_j^T(t)\mathrm{d}t D_j^T \tilde{F}_j(d_j \otimes I_k).$$

Due to the definition of D_j in (3.6) and $\tilde{B}_j = (d_j \otimes I_r)B'_j$, the above equation becomes

$$S_{jj} = \underbrace{\int_0^{t_f} B'_j(t)R_j^{-1}(t)(B'_j)^T(t)\mathrm{d}t}_{a} \underbrace{(d_j^T \otimes I_r)\tilde{F}_j(d_j \otimes I_k)}_{b}.$$

Since both term a and term b are positive semi-definite, all the eigenvalues of matrix S_{jj} have nonnegative real parts. $\qquad\square$

Lemma 3.2. *If $(-M)$ is Hurwitz for matrix M defined in (3.14a), supposing that matrix P is the unique positive definite solution to the Lyapunov equation*

$$-M^T P - PM = -I, \tag{3.24}$$

then for linear system

$$\dot{\theta} = -gM\theta \tag{3.25}$$

with the initial state $\theta(0)$, the inequalities

$$\int_0^{t_f} \|\theta\|_2^2 dt \leq \frac{2\gamma_{\max}V(0)}{g\gamma_{\min}}(1 - e^{-g\gamma_{\min}t_f}), \tag{3.26a}$$

$$\int_0^{t_f} \|\theta\|_2 dt \leq \frac{2\sqrt{2\gamma_{\max}V(0)}}{g\gamma_{\min}}(1 - e^{-g\gamma_{\min}t_f/2}). \tag{3.26b}$$

hold, where $\|\cdot\|_2$ stand for Euclidean norm, $V(0) = 1/2\theta^T(0)P\theta(0)$, $\gamma_{\min} = 1/\lambda_{\max}(P)$, $\lambda_{\max}(P)$ is the largest eigenvalue of matrix P, $\gamma_{\max} = 1/\lambda_{\min}(P)$, and $\lambda_{\min}(P)$ is the smallest eigenvalue of matrix P.

Proof. If matrix $(-M)$ is Hurwitz, we consider the quadratic Lyapunov function $V = 1/2\theta^T P\theta$ for system (3.25), where matrix P is the solution to the Lyapunov equation (3.24). The derivative

of the Lyapunov function along the trajectory of system (3.25) is

$$\dot{V} = \frac{1}{2}\theta^T(-gM^TP - gPM)\theta = -\frac{g}{2}\|\theta\|_2^2 \leq -\frac{g}{2\lambda_{\max}(P)}\theta^TP\theta = -g\gamma_{\min}V$$

where the last inequality is due to the property $\theta^TP\theta \leq \lambda_{\max}(P)\|\theta\|_2^2$. The above differential inequality yields

$$\frac{1}{2}\lambda_{\min}(P)\|\theta\|_2^2 \leq \frac{1}{2}\theta^TP\theta = V(t) \leq e^{-g\gamma_{\min}t}V(0) \implies \|\theta\|_2^2 \leq 2\gamma_{\max}e^{-g\gamma_{\min}t}V(0).$$

Therefore, integrating the above inequality from 0 to t_f yields (3.26a) and the square root of the above inequality from 0 to t_f yields (3.26b). $\qquad\square$

We now present the ε-Nash equilibrium for the proposed strategy (3.22) as the following theorem:

Theorem 3.5. *For the differential game in an N-agent system under distributed information defined by system dynamics in (3.11) and performance indices in (5.9), if $(-M)$ is Hurwitz for matrix M defined in (3.14a), then the online computing strategies described by (3.22) form an ε-Nash equilibrium where*

$$\varepsilon = \max_{i=1,\cdots,N} \varepsilon_i \tag{3.27}$$

where

$$\varepsilon_i = \frac{2\|\tilde{F}_{Ri}\|_2\gamma_{\max}V(0)}{g\gamma_{\min}}(1 - e^{-g\gamma_{\min}t_f}) + \frac{8\|\tilde{M}_i\|_2W_{\max}^2\gamma_{\max}V(0)}{g^2\gamma_{\min}^2}(1 - e^{-g\gamma_{\min}t_f/2})^2, \tag{3.28}$$

scalars $\gamma_{\min}$, $\gamma_{\max}$, *and* $V(0)$ *are defined in Lemma 3.2 with* $\theta(0) = z^f(0) - M^{-1}z(0)$,

$$\tilde{F}_{Ri} = D_i^T \tilde{F}_i D_i \tilde{B}_i R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i, \tag{3.29a}$$

$$\tilde{M}_i = M_{Fi} + t_f M_i^T \tilde{F}_{Ri} M_i, \tag{3.29b}$$

$$M_{Fi} = M^T(d_i \otimes I_r)(I + S_{ii})^{-1}\tilde{F}_{ii}(I + S_{ii})^{-1}(d_i^T \otimes I_r)M, \tag{3.29c}$$

$$\tilde{F}_{ii} = (d_i^T \otimes I_r)\tilde{F}_i(d_i \otimes I_r), \tag{3.29d}$$

$$M_i = I - (d_i \otimes I_r)(I + S_{ii})^{-1}(d_i^T \otimes I_r)M, \tag{3.29e}$$

$$W_{\max} = \max_{0 \le t \le t_f} \|W(t)\|_2, \quad W(t) = \sum_{j=1}^N \tilde{B}_j R_j^{-1} \tilde{B}_j^T D_j^T \tilde{F}_j D_j, \tag{3.29f}$$

and matrix S_{ii} *is defined in (3.23).*

Proof. First of all, if knowing every other agent will choose the online computing strategy (3.22), the best strategy of agent i in response to these strategies is

$$u_i^* = - R_i^{-1}\tilde{B}_i^T D_i^T \tilde{F}_i D_i z^*(t_f) \tag{3.30}$$

which is obtained from (3.16), where z_i^* is agent i's state trajectory under the best reaction strategy u_i^* in (3.30). Denoting z_j as the state trajectory of agent j ($j \neq i$) for all $e_{ij} \in \mathcal{E}$ under strategy (3.22) yields

$$D_i z^*(t_f) = (d_i \otimes I_r)z_i^*(t_f) + \sum_{e_{ij} \in \mathcal{E}, j \neq i} (d_j \otimes I_r)z_j(t_f).$$

Then, u_i^* in (3.30) becomes

$$u_i^* = - R_i^{-1}\tilde{B}_i^T D_i^T \tilde{F}_i \left[(d_i \otimes I_r)z_i^*(t_f) + \sum_{e_{ij} \in \mathcal{E}, i \neq j} (d_j \otimes I_r)z_j(t_f) \right]$$

$$= - R_i^{-1}\tilde{B}_i^T D_i^T \tilde{F}_i [D_i z(t_f) - (d_i \otimes I_r)\Delta z_i(t_f)] \tag{3.31}$$

35

where $\Delta z_i = z_i - z_i^*$. We denote J_i as the performance index value of (3.8) if every agent chooses strategy (3.22) and J_i^* as the performance index value of (3.8) if agent i chooses strategy (3.31) and every other agent chooses strategy (3.22). The difference between J_i and J_i^* is

$$
\begin{aligned}
J_i - J_i^* =& \frac{1}{2}(\|D_i z(t_f)\|_{\tilde{F}_i}^2 - \|D_i z^*(t_f)\|_{\tilde{F}_i}^2) + \frac{1}{2}\int_0^{t_f}(\|u_i\|_{R_i}^2 - \|u_i^*\|_{R_i}^2)\mathrm{d}t \\
=& \frac{1}{2}[D_i z(t_f) - D_i z^*(t_f)]^T \tilde{F}_i[D_i z(t_f) + D_i z^*(t_f)] + \frac{1}{2}\int_0^{t_f}(u_i - u_i^*)_i^T R_i(u_i + u_i^*)\mathrm{d}t \\
=& \frac{1}{2}\Delta z_i^T(t_f)\tilde{F}_{ii}\Delta z_i(t_f) + \frac{1}{2}\int_0^{t_f}\Delta u_i^T R_i(\Delta u_i + 2u_i^*)\mathrm{d}t
\end{aligned}
\tag{3.32}
$$

where $\tilde{F}_{ii}$ is defined in (3.29d) and $\Delta u_i = u_i - u_i^*$. Since the dynamics of Δz_i is as follows:

$$
\Delta \dot{z}_i = \dot{z}_i - \dot{z}_i^* = B_i' u_i - B_i' u_i^* = B_i' \Delta u_i,
\tag{3.33}
$$

integrating the above equation from 0 to t_f yields

$$
\Delta z_i(t_f) = \int_0^{t_f} B_i' \Delta u_i \mathrm{d}t
$$

Substituting the above equation into (3.32) and recalling the expression of u_i^* in (3.30) yield

$$
J_i - J_i^* = \frac{1}{2}\Delta z_i^T(t_f)\tilde{F}_{ii}\Delta z_i(t_f) + \frac{1}{2}\int_0^{t_f}\Delta u_i^T R_i \Delta u_i \mathrm{d}t.
\tag{3.34}
$$

Clearly, the value of $(J_i - J_i^*)$ shown above is always nonnegative which is as expected. To find the value or upper bound of $(J_i - J_i^*)$, it is necessary to find the values of $\Delta z_i(t_f)$ and Δu_i. Toward that end, first of all, solving the differential equation in (3.21) yields

$$
z^f(t) = M^{-1}z(0) + \theta(t).
\tag{3.35}
$$

where $\theta(t)$ is defined in (3.25) with the initial state $\theta(0) = z^f(0) - M^{-1}z(0)$. The system dynamics

36

(3.11) when every agent chooses strategy (3.22b) becomes

$$\dot{z} = -\sum_{j=1}^{N} \tilde{B}_j R_j^{-1} \tilde{B}_j^T D_i^T \tilde{F}_j D_j z^f \triangleq -W z^f \tag{3.36}$$

where W is defined in (3.29f). Substituting (3.35) into (3.36) yields

$$\dot{z} = -WM^{-1}z(0) - W\theta. \tag{3.37}$$

Since

$$\int_0^{t_f} W \mathrm{d}t = \sum_{j=1}^{N} S_j F_j$$

where matrices F_j and S_j are defined in (3.7) and (3.14b), integrating equation (3.37) from 0 to t_f yields

$$z(t_f) = \left(I - \sum_{j=1}^{N} S_j F_j M^{-1} \right) z(0) - \int_0^{t_f} W\theta \mathrm{d}t. \tag{3.38}$$

Recalling the definition of matrix M in (3.14a), we have

$$\left(I + \sum_{j=1}^{N} S_j F_j \right) M^{-1} = I \quad \Longrightarrow \quad M^{-1} = I - \sum_{j=1}^{N} S_j F_j M^{-1}.$$

Hence, equation (3.38) becomes

$$z(t_f) = M^{-1}z(0) - \int_0^{t_f} W\theta \mathrm{d}t. \tag{3.39}$$

Second, substituting u_i^* in (3.31) into (3.10) yields

$$\dot{z}_i^* = -(d_i^T \otimes I_r)\tilde{B}_i R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i \left[D_i z(t_f) - (d_i \otimes I_r)\Delta z_i(t_f) \right]$$

37

Integrating the above equation from 0 to t_f yields

$$z_i^*(t_f) = z_i(0) - (d_i^T \otimes I_r) S_i D_i^T \tilde{F}_i \left[D_i z_i(t_f) - (d_i \otimes I_r) \Delta z_i(t_f) \right]$$
$$= z_i(0) + S_{ii} \Delta z_i(t_f) - (d_i^T \otimes I_r) S_i F_i z(t_f)$$

where S_{ii} is as defined in (3.23). Therefore, substituting the above equation into the expression of $\Delta z_i(t_f)$ yields

$$\Delta z_i(t_f) = z_i(t_f) - z_i^*(t_f) = z_i(t_f) - z_i(0) - S_{ii} \Delta z_i(t_f) + (d_i^T \otimes I_r) S_i F_i z(t_f)$$

After some manipulations, we arrive at

$$(I + S_{ii}) \Delta z_i(t_f) = (d_i^T \otimes I_n)[M z(t_f) - z(0)]$$

As we showed in Lemma 3.1, all the eigenvalues of S_{ii} in the right hand side of the above equation have nonnegative real parts and hence matrix $(I + S_{ii})$ is invertible. Therefore,

$$\Delta z_i(t_f) = (I + S_{ii})^{-1} (d_i^T \otimes I_r) \left[M z(t_f) - z(0) \right] \tag{3.40}$$

Therefore, substituting (3.39) into (3.40) yields the value of $\Delta z_i(t_f)$ as follows:

$$\Delta z_i(t_f) = - (I + S_{ii})^{-1} (d_i^T \otimes I_r) M \int_0^{t_f} W \theta \mathrm{d}t. \tag{3.41}$$

The value of Δu_i obtained as follows:

$$
\begin{aligned}
\Delta u_i = u_i - u_i^* = & - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^f + R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^*(t_f) \\
= & - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^f + R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z(t_f) - R_i^{-1} \tilde{B}_i D_i^T \tilde{F}_i (d_i \otimes I_r) \Delta z_i(t_f) \\
= & - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i [z^f - z(t_f)] - R_i^{-1} \tilde{B}_i D_i^T \tilde{F}_i (d_i \otimes I_r) \Delta z_i(t_f), && (3.42)
\end{aligned}
$$

Recalling (3.18) and (3.39), the value of $[z^f - z(t_f)]$ in the above equation can be obtained as

$$
z^f - z(t_f) = \theta + \int_0^{t_f} W \theta \mathrm{d}t \qquad (3.43)
$$

Substituting (3.41) and (3.43) into (3.42) yields

$$
\begin{aligned}
\Delta u_i = & - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i \left(\theta + M_i \int_0^{t_f} W \theta \mathrm{d}t \right) \\
\triangleq & - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i v_i, \qquad (3.44)
\end{aligned}
$$

where $v_i = \theta + M_i \int_0^{t_f} W \theta \mathrm{d}t$ and matrix M_i is defined in (3.29e). Given the value of $\Delta z_i(t_f)$ in (3.41) and Δu_i in (3.44), the upper bound of $(J_i - J_i^*)$ can be derived from (3.34) as follows:

$$
\begin{aligned}
J_i - J_i^* = & \frac{1}{2} \Delta z_i^T(t_f) \tilde{F}_{ii} \Delta z_i(t_f) + \frac{1}{2} \int_0^{t_f} \Delta u_i^T R_i \Delta u_i \mathrm{d}t \\
= & \frac{1}{2} \left(\int_0^{t_f} W \theta \mathrm{d}t \right)^T M_{Fi} \left(\int_0^{t_f} W \theta \mathrm{d}t \right) + \frac{1}{2} \int_0^{t_f} v_i^T \tilde{F}_{Ri} v_i \mathrm{d}t \qquad (3.45)
\end{aligned}
$$

where matrix M_{Fi} is defined in (3.29c) and matrix $\tilde{F}_{Ri}$ is defined in (3.29a). Since

$$
\begin{aligned}
\frac{1}{2} \int_0^{t_f} v_i^T \tilde{F}_{Ri} v_i \mathrm{d}t = & \frac{1}{2} \int_0^{t_f} \left(\theta + M_i \int_0^{t_f} W \theta \mathrm{d}t \right)^T \tilde{F}_{Ri} \left(\theta + M_i \int_0^{t_f} W \theta \mathrm{d}t \right) \mathrm{d}t \\
\leq & \int_0^{t_f} \theta^T \tilde{F}_{Ri} \theta \mathrm{d}t + t_f \left(\int_0^{t_f} W \theta \mathrm{d}t \right)^T M_i^T \tilde{F}_{Ri} M_i \left(\int_0^{t_f} W \theta \mathrm{d}t \right) \qquad (3.46)
\end{aligned}
$$

39

Substituting (3.46) into (3.45) yields

$$J_i - J_i^* \leq \int_0^{t_f} \theta^T \tilde{F}_{Ri} \theta \mathrm{d}t + \left(\int_0^{t_f} W \theta \mathrm{d}t \right)^T \tilde{M}_i \left(\int_0^{t_f} W \theta \mathrm{d}t \right)$$

$$\leq \|\tilde{F}_{Ri}\|_2 \int_0^{t_f} \|\theta\|_2^2 \mathrm{d}t + \|\tilde{M}_i\|_2 W_{\max}^2 \left(\int_0^{t_f} \|\theta\|_2 \mathrm{d}t \right)^2$$

where matrix $\tilde{M}_i$ is defined in (3.29b) and $W_{\max}$ is defined in (3.29f). Recalling Lemma 3.2, substituting inequalities in (3.26) into the above inequality yields (3.28). The maximum value in $\{\varepsilon_1, \cdots, \varepsilon_N\}$ will satisfy the inequalities (2.7). Therefore, online computing strategies (3.22) form an ε-Nash equilibrium. □

Note that as shown in (3.28), it is clear that the value of ε_i decreases as g becomes larger. Therefore, we can claim that there exists a scalar g such that the online computing strategies in (3.22) forms an ε-Nash equilibrium that can be arbitrarily close to the Nash equilibrium in (3.13).

3.2.2   Noninferior Strategy Design

The open-loop Noninferior solution can also be derived using Pontryagin's minimum principle and is presented as follows.

Theorem 3.6. *For the differential game in an N-agent system under distributed information defined by system dynamics in (3.11) and performance indices in (5.9), the strategies*

$$u_i = -\frac{1}{\alpha_i} R_i^{-1} \tilde{B}_i^T \sum_{j=1}^N \alpha_j D_j^T \tilde{F}_j D_j M_p^{-1} z(0) \qquad \forall i = 1, \cdots, N \tag{3.47}$$

form an open-loop noninferior solution where $0 \leq \alpha_j \leq 1$ for all $j = 1, \cdots, N$ and $\sum_{j=1}^N \alpha_j = 1$

if matrix M_P defined by

$$M_P = \left(I_{Nr} + \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\alpha_k}{\alpha_j} S_j D_k^T \tilde{F}_k D_k \right), \tag{3.48}$$

is invertible, where matrix S_j is defined in (3.14b).

Proof. Given system dynamics in (3.11) and performance indices in (5.9), to find the noninferior solution, we define a convex combination of $J_1, \cdots, J_N$ as shown in (3.3), which is

$$J = \frac{1}{2} z^T(t_f) \left(\sum_{j=1}^{N} \alpha_j D_j^T \tilde{F} D_j \right) z(t_f) + \frac{1}{2} \int_0^{t_f} \left(\sum_{j=1}^{N} \alpha_j u_j^T R_j u_j \right) \mathrm{d}t. \tag{3.49}$$

We define the Hamiltonian as

$$H = \frac{1}{2} \left(\sum_{j=1}^{N} \alpha_j u_j^T R_j u_j \right) + \lambda^T \sum_{j=1}^{N} \tilde{B}_j u_j$$

where vector $\lambda \in \mathbb{R}^{Nr}$ is the Lagrangian multiplier. Since the second order partial derivative of H with respect to $u_1, \cdots, u_N$ are all equal to $\alpha_i R_i$ and hence is positive define, the following conditions for $u_1, \cdots, u_N$ to minimize the performance index are necessary and sufficient

$$\dot{z} = \frac{\partial H}{\partial \lambda} = \sum_{j=1}^{N} \tilde{B}_j u_j, \tag{3.50a}$$

$$\dot{\lambda} = -\frac{\partial H}{\partial z} = 0, \quad \lambda(t_f) = \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j z(t_f), \tag{3.50b}$$

$$\frac{\partial H}{\partial u_i} = \alpha_i R_i u_i + \tilde{B}_i^T \lambda = 0 \quad \forall i = 1, \cdots, N. \tag{3.50c}$$

41

Condition (3.50b) indicates that λ is a constant vector and hence

$$\lambda(t) = \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_i z(t_f) \quad \forall t \in [0, t_f]$$

From (3.50c), we obtain

$$u_i = -\frac{1}{\alpha_i} R_i^{-1} \tilde{B}_i^T \lambda = -\frac{1}{\alpha_i} R_i^{-1} \tilde{B}_i^T \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j z(t_f). \tag{3.51}$$

Substituting (3.59) into (3.50a) and integrating both sides from 0 to t_f yield

$$\left(I_{Nr} + \sum_{j=1}^{N} \frac{1}{\alpha_j} S_j \sum_{k=1}^{N} \alpha_k D_k^T \tilde{F}_k D_k \right) z(t_f) = z(0)$$

$$M_P z(t_f) = z(0) \tag{3.52}$$

where matrix M_P is defined in (3.48) and matrix S_j is defined in (3.14b). Therefore, if the matrix M_P is invertible, then

$$z(t_f) = M_P^{-1} z(0). \tag{3.53}$$

Substituting (3.53) into (3.59) yields the open-loop Nash strategy (3.47). $\qquad\square$

Again, note that M_P^{-1} in (3.47) is generally a full matrix. Therefore, implementing open-loop noninferior strategy u_i in (3.47) will still require every agent to have complete knowledge of the initial state of all the agents, $z(0) = [z_1^T(0), \cdots, z_N^T(0)]^T$. However, just like the proposed open-loop Nash strategy design approach in the previous section, we are also able to apply the same idea to the open-loop noninferior strategy design. Toward that end, we have the following result:

Theorem 3.7. *For the differential game in an N-agent system environment defined by system dynamics (3.11) and performance indices (5.9), if*

1. *matrix $(-M_P)$ in (3.48) is Hurwitz and*

2. *agent i updates the estimate of its state, $z_i^f \in \mathbb{R}^r$, according to*

$$\dot{z}_i^f = g\left[z_i(0) - z_i^f - (d_i^T \otimes I_r)\frac{1}{\alpha_i}S_i \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j z^f\right] \quad \forall i = 1, \cdots, N \qquad (3.54)$$

from any initial condition $z_i^f(0)$, where g is a positive scalar and $z^f = [(z_1^f)^T \ \cdots \ (z_N^f)^T]^T$,

then

$$\lim_{t \to \infty} z_i^f(t) = (d_i^T \otimes I_r)M_p^{-1}z(0). \qquad (3.55)$$

Proof. The proof is in the same fashion as the one for theorem 3.4. Stacking equation (3.54) from $i = 1$ to $i = N$ yields

$$\begin{aligned}
\dot{z}^f &= g\left[z(0) - z^f - \sum_{j=1}^{N}\frac{1}{\alpha_j}S_j \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j z^f\right] \\
&= g\left[z(0) - M_P z^f\right].
\end{aligned} \qquad (3.56)$$

where matrix M_P is defined in (3.48). If matrix $(-M_P)$ is Hurwitz, linear system (3.56) with respect to z^f is asymptotically stable starting from any initial condition $z^f(0)$ and z^f asymptotically converges to the equilibrium of the differential equation (3.56), that is,

$$\lim_{t \to \infty} z^f(t) = M_P^{-1}z(0).$$

Multiplying $(d_i^T \otimes I_r)$ on both sides of the above equation yields (3.55). $\qquad \square$

Note that one can increase the positive scalar g to achieve satisfactory convergence speed of the state estimation process. Also note that it is not clear so far whether implementing (3.54) only re-

quires each agent to have the information available to it only. In fact, there exist several conditions on the information graph and agents' performance indices such that the estimation law (3.54) can be carried out in a distributed manner. These conditions are presented in the following corollary.

Corollary 3.1. *For the differential game in an N-agent system environment defined by system dynamics (3.11) and performance indices (5.9), if*

1. *the information graph among the agents is undirected[1] and*

2. *matrix $\tilde{F}_j$ has the following structure*

$$\tilde{F}_j = (d_j d_j^T) \otimes \tilde{F}_j^{jj} + \sum_{e_{jk} \in \mathcal{E}, k \neq j} [(d_k d_k^T) \otimes \tilde{F}_j^{kk} + (d_j d_k^T) \otimes \tilde{F}_j^{jk} + (d_k d_j^T) \otimes \tilde{F}_j^{kj}], \quad (3.57)$$

for all $j = 1, \cdots, N$,

then the state estimation law in (3.54) can be expressed as

$$\dot{z}_i^f = g \left\{ z_i(0) - z_i^f - S_i' \left[\sum_{e_{ij} \in \mathcal{E}} \tilde{F}_i^{ij} z_i^f + \sum_{e_{ji} \in \mathcal{E}, j \neq i} \frac{\alpha_j}{\alpha_i} \left(\tilde{F}_j^{ij} z_j^f + \tilde{F}_j^{ii} z_i^f \right) \right] \right\}. \quad (3.58)$$

for all $i = 1, \cdots, N$.

Proof. For agent i, matrix D_j defined in (3.6) can be expressed as

$$D_j = \begin{cases} (d_i d_i^T) \otimes I_r + \displaystyle\sum_{e_{jk} \in \mathcal{E}, k \neq i} (d_k d_k^T) \otimes I_r & \text{if } e_{ji} \in \mathcal{E} \\ \displaystyle\sum_{e_{jk} \in \mathcal{E}, k \neq i} (d_k d_k^T) \otimes I_r & \text{if } e_{ji} \notin \mathcal{E} \end{cases}$$

---

[1] A graph is undirected if every edge is bidirectional, that is, edges $e_{ij} \in \mathcal{E}$ indicates $e_{ji} \in \mathcal{E}$.

for all $j = 1, \cdots, N$ and hence recalling (3.14b), we have

$$S_i \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j = (d_i \otimes I_r) S_i' \sum_{e_{ji} \in \mathcal{E}} \alpha_j (d_i^T \otimes I_r) \tilde{F}_j D_j, \qquad (3.59)$$

where $S_i' = \int_0^{t_f} B_i' R^{-1} (B_i')^T dt$. Substituting (3.57) into (3.59) yields

$$\begin{aligned}
S_i \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j = (d_i \otimes I_r) S_i' \sum_{e_{ji} \in \mathcal{E}} \alpha_j (d_i^T \otimes I_r) \Bigg\{ (d_k d_k^T) \otimes \tilde{F}_j^{jj} \\
+ \sum_{e_{jk} \in \mathcal{E}, k \neq j} [(d_k d_k^T) \otimes \tilde{F}_j^{kk} + (d_j d_k^T) \otimes \tilde{F}_j^{jk} + (d_k d_j^T) \otimes \tilde{F}_j^{kj}] \Bigg\} D_j
\end{aligned}$$

After some mathematical manipulations to the above equation and using the property of a undirected graph (bidirectional edges), we arrive at

$$S_i \sum_{j=1}^{N} \alpha_j D_j^T \tilde{F}_j D_j = (d_i \otimes I_r) S_i' \left[\alpha_i \sum_{e_{ij} \in \mathcal{E}} (d_j^T \otimes \tilde{F}_i^{ij}) + \sum_{e_{ij} \in \mathcal{E}, j \neq i} \alpha_j (d_j^T \otimes \tilde{F}_j^{ij} + d_i^T \otimes \tilde{F}_j^{ii}) \right]$$
$$(3.60)$$

Therefore, substituting (3.60) into (3.54) yields (3.58) $\qquad \square$

It is clear that under the conditions in Corollary 3.1, the state estimation law expressed in (3.57) for agent i only requires it to receive the state estimate z_j^f from agent j for all $e_{ij} \in \mathcal{E}$. Therefore, this estimation law can be carried out by individual agents in a distributed manner. Note that the two conditions in Corollary 3.1 can be successfully satisfied for many applications of the multi-agent system and this will be illustrated in an example later in Chapter 5. Compared with the state estimation law for the Nash equilibrium in the previous section (which does not require each agent to have the information of the system dynamics and performance indices of other agents and the overall information topology), the proposed state estimation law for the noninferior solution does not require each agent to know the overall information topology either, however, it requires each

agent to have the information of the convex combination parameters $\alpha_1, \cdots, \alpha_N$ (which need be assigned to individual agents due to the nature of the cooperative game) and the coefficients $\tilde{F}_j^{ij}$ and $\tilde{F}_j^{ii}$ in the performance index of agent j for $e_{ij \in \mathcal{E}}$.

Similar to the open-loop Nash strategy design algorithm, to implement the open-loop noninferior strategy, the state estimation law (3.58) can be carried out offline first until a certain convergent value of the state is achieved. Moreover, an online open-loop noninferior strategy design algorithm based on (3.59) can be proposed similar to (3.22) as follows:.

$$\dot{z}_i^f = g \left\{ z_i(0) - z_i^f - S_i' \left[\sum_{e_{ij} \in \mathcal{E}} \tilde{F}_i^{ij} z_i^f + \sum_{e_{ji} \in \mathcal{E}, j \neq i} \frac{\alpha_j}{\alpha_i} \left(\tilde{F}_j^{ij} z_j^f + \tilde{F}_j^{ii} z_i^f \right) \right] \right\} \tag{3.61a}$$

$$u_i = -R_i^{-1} (B_i')^T \left[\sum_{e_{ij} \in \mathcal{E}} (\tilde{F}_i^{ij} z_i^f) + \sum_{e_{ij} \in \mathcal{E}, j \neq i} \frac{\alpha_j}{\alpha_i} (\tilde{F}_j^{ij} z_j^f + \tilde{F}_j^{ii} z_i^f) \right] \tag{3.61b}$$

for all $i = 1, \cdots, N$. Although the above online computing algorithm is approximate to the actual noninferior strategy, it is obvious that the convergence speed of (3.61a) can be made arbitrarily fast by adjusting the value of g and hence the strategy in (3.61) can be arbitrarily close to the noninferior strategy by choosing a proper value of g.

3.3 Extensions

In the previous sections, we have proposed the distributed open-loop Nash strategy design approach and distributed open-loop noninferior strategy design approach for differential games in the multi-agent system under distributed information. Several extensions are presented as follows:

1. Note that these strategies are developed based on performance indices (5.9) which do not contain the integral costs on the quadratic form of the state, while the general performance

indices in (2.5) contain such costs. The advantage of including such integral costs on the state is to improve the transient behavior of the state trajectory under the resulting Nash strategies or noninferior strategies. However, in order to develop a distributed game design approach based on the novel state estimation law, it is required that such integral costs on the state to be eliminated. Fortunately, for most of the applications in the multi-agent systems, such as rendezvous, formation control, flocking, etc., a satisfactory performance on the terminal state is much more important than the performance on the transient state. Therefore, adopting the performance indices (5.9) is a reasonable choice in most of the applications and hence the proposed approaches are applicable. Alternatively, one possible approach to improve the transient behavior of the state is dividing the entire game horizon $[0, t_f]$ into l intervals, that is,

$$[0, t_f] = [0, t_1] \cup (t_1, t_2] \cup \cdots \cup (t_{l-2}, t_{l-1}] \cup (t_{l-1}, t_f].$$

The proposed distributed game strategy design approaches can then be carried out by the agents at each and every time interval based on the performance indices (5.9) with the entire time horizon $[0, t_f]$ replaced by the above smaller time intervals and the terminal state $z(t_f)$ replaced by $z(i)$ for all $i = t_1, \cdots, t_{l-1}, t_f$. As an approximate of the original problem, this alternative game problem with several smaller horizons will not render game strategies that are exactly the same as the ones derived based on the original performance indices. However, since this approach takes several sampled states at the transient time instants $t_1, t_2, \cdots, t_{l-2}, t_{l-1}$ into account, the resulting state trajectory will have better transient performance than the one over only one horizon $[0, t_f]$.

2. For the differential game over the infinite time horizon, that is, $t_f = \infty$, in order to implement the proposed open-loop distributed strategy design approach, one possible approach is to combine this approach with the receding horizon control technique [72] which has been widely used and given very good results in practice applications. Specifically, in this case,

47

we will consider the following receding horizon performance indices instead of (5.9)

$$J_i = \frac{1}{2}\|z(\tau + t_f)\|_{F_i}^2 + \frac{1}{2}\int_\tau^{\tau+t_f} \|u_i(t)\|_{R_i(t)}^2 \mathrm{d}t \quad \forall i = 1, \cdots, N. \tag{3.62}$$

The mechanism of implementing the receding horizon control is shown in Figure. 3.1.
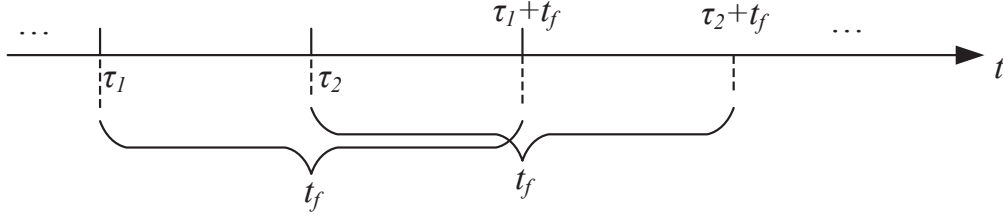

Figure 3.1: Receding horizon control mechanism

Specifically, at time $\tau = \tau_1$, the Nash equilibrium or noninferior solution of the differential game with performance indices (3.62) is solved for the time interval $[\tau_1, \tau_1 + t_f]$. The agents will implement the corresponding Nash strategies or noninferior strategies from τ_1 to τ_2 where $\tau_2 \leq \tau_1 + t_f$. The Nash equilibrium and noninferior solution will be recalculated from τ_2 to $\tau_2 + t_f$ and the agent will implement the corresponding strategies from τ_2 to τ_3. This procedures will be repeated as the game proceeds. We should point out that since the game is over the infinite time horizon, the stability of the multi-agent system is fairly important, however, this issue with the receding horizon control technique is largely open and still under investigation.

3. As we know, the feedback game strategy is preferred than the open-loop one in many real life applications because the feedback strategies can react to the instantaneous disturbance in the states. A practical way to convert the proposed open-loop design approach into a feedback-like type is to utilize the sampled-Nash approach [73]. Toward that end, we consider the

following performance indices instead of (5.9):

$$J_i = \frac{1}{2}\|z(t_f)\|_{F_i}^2 + \frac{1}{2}\int_{t_k}^{t_f}\|u_i(t)\|_{R_i(t)}^2 \mathrm{d}t \quad \forall i = 1, \cdots, N, t = 1, \cdots, l. \tag{3.63}$$

where $0 = t_1 \leq t_2 \leq \cdots \leq t_l < t_f$. Combining the distributed online computing Nash strategy in (3.22), the sampled distributed open-loop Nash strategy design algorithm is proposed as follows.

Algorithm 3.1. *At $t = t_k$ for all $j = 1, \cdots, l$,*

1. *Agents measure and calculate $z_1(t_k), \cdots, z_N(t_k)$.*

2. *Agent implement (3.22) (for Nash strategy) or (3.61) (for noninferior strategy) with $z_i(0)$ replaced by $z_i(t_k)$ for $i = 1, \cdots, N$, respectively.*

3. *Once $t = t_{k+1}$ arrives, the agents repeat the step 1 and 2 by letting $t_k \to t_{k+1}$.*

Clearly, since the agents measure the states multiple times during the process and will hence be more aware of the unexpected change in the system.

In this chapter, we considered the open-loop game strategy design approach in the multi-agent system under distributed information structure. The basic idea of the proposed approach is to let the agents in the system exchange certain information among themselves according to the information graph such that their strategies asymptotically converge to the Nash or noninferior strategies that can only be implemented under global information originally. This approach can be applied to most of the applications in multi-agent systems and can also be extended to the differential games over the infinite time horizon and differential games under feedback information structure.

CHAPTER 4: FEEDBACK GAME STRATEGIES

In this chapter, we consider the game strategy design under distributed feedback information structure. The Riccati equation approach and the conventional optimal output feedback design approach are introduced first. A novel distributed game strategy design approach is then proposed based on the concept of best achievable performance indices.

4.1 Riccati Equation Approach

Feedback Nash Equilibrium Strategy: The feedback Nash equilibrium can be obtained by either using the Pontryagin's minimum principle or solving the Hamilton-Jacobi-Bellman partial differential equations. The well-known feedback linear quadratic Nash equilibrium [7] is presented as the following theorem without proof.

Theorem 4.1. *For the differential game in an N-agent system environment defined by system dynamics (2.2) and performance indices (2.5), the strategies*

$$u_i = -R_i^{-1}(t)B_i^T(t)P_i(t)x(t) \quad \forall i = 1, \cdots, N \tag{4.1}$$

form a feedback Nash equilibrium, where matrix $P_i(t)$ is the solution to the following coupled differential Riccati equations

$$\dot{P}_i + P_i B_i R_i^{-1} B_i^T P_i + P_i \bar{A} + \bar{A}^T P_i + C_i^T Q_i C_i = 0 \quad \forall i = 1, \cdots, N \tag{4.2}$$

with the boundary condition $P_i(t_f) = C_i^T(t_f)F_i C_i(t_f)$ and $\bar{A} = A - \sum_{j=1}^{N} B_j R_j^{-1} B_j^T P_j$.

It is clear that since the coupled differential Riccati equation for the open-loop Nash equilibrium

in (3.2) and the ones for the feedback Nash equilibrium in (4.2) are different, the open-loop Nash strategy in (3.1) and feedback one in (4.1) are completely different.

Feedback Noninferior Solution Strategy: Since the noninferior solution is derived by solving the linear quadratic optimal control problem with respect to the convex combination of the performance indices in (3.3), the noninferior solution under the feedback information structure will turn out to be in the same as (3.4) with the product of $[\phi(t,0)x(0)]$ replaced with $x(t)$. We present the following theorem.

Theorem 4.2. *For the differential game in an N-agent system environment defined by system dynamics in (2.2) and performance indices in (2.5), the strategies*

$$u_i(t) = -\frac{1}{\alpha_i}R_i^{-1}(t)B_i^T(t)P(t)x(t) \quad \forall i = 1, \cdots, N \tag{4.3}$$

form a feedback noninferior solution, where matrix $P(t)$ is the solution to the differential Riccati equation (3.5) with the boundary condition $P(t_f) = \sum_{j=1}^{N}\alpha_j C_j^T(t_f)F_j C_j(t_f)$.

To derive the feedback Nash equilibrium or feedback noninferior solution, one has to solve the differential Riccati equations (4.2) or differential Riccati equation (3.5) backward in time. Unfortunately, like the open-loop Nash equilibrium and noninferior solution, the solution will generally become a full matrix. Therefore, there is generally no way for the agents in the system to implement the feedback Nash strategy or noninferior strategy without the complete information of the state information, $x(t)$, at every instant of time t. To overcome this issue, we introduce an existing distributed game strategy design approach based on the optimal output feedback control in the next section.

4.2 Optimal Output Feedback Approach

In this section, we consider the differential game in the N-agent system under distributed feedback information structure. In order to conform to the underlying information graph constraint, the agents' linear structured strategies are in the following form:

$$u_i^s = \sum_{e_{ij} \in \mathcal{E}} K_{ij} y_j \triangleq K_i^s D_i y = K_i^s D_i C x \quad \forall i = 1, \cdots, N,$$

where the superscript s in u_i^s means that the strategy is structured, $K_i^s = [K_{i1}^s \ \cdots \ K_{iN}^s] \in \mathbb{R}^{m_i \times Nr}$, and matrix D_i is defined in (3.6). The structure of u_i^s indicates that the strategy of each agent can only be a feedback of the output information that is available to it only according to the information graph. Denoting $\hat{C}_i = D_i C$, the structured strategies are expressed as

$$u_i^s = K_i^s \hat{C}_i x \quad \forall i = 1, \cdots, N. \tag{4.4}$$

which are apparently in the output feedback form. The basic idea of applying the optimal output feedback approach [54] to the game strategy design is to parameterizing matrix K_i^s in (4.4) and derive the Nash equilibrium or noninferior solution directly with respect to this variables $K_1^s, \cdots, K_N^s$. In what follows, we present the distributed Nash strategy and noninferior strategy design using the optimal output feedback approach.

4.2.1 Nash Strategy Design

With parameterized $K_1^s, \cdots, K_N^s$ in (4.4), the Nash equilibrium is presented as the following theorem based on the optimal output feedback approach.

Theorem 4.3. *For the differential game in an N-agent system under distributed feedback infor-*

mation structure defined by system dynamics (2.2) and performance indices (2.5), the strategies in

(4.4) form a feedback Nash equilibrium if the following equations holds:

$$\dot{x} = \bar{A}x, \tag{4.5a}$$

$$\bar{A} = \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right), \tag{4.5b}$$

$$\dot{P}_i + P_i \bar{A} + \bar{A}^T P_i + C^T Q_i C + \hat{C}_i^T (K_i^s)^T R_i K_i^s \hat{C}_i = 0 \tag{4.5c}$$

$$P_i(t_f) = C_i^T F_i C \tag{4.5d}$$

$$R_i K_i^s \hat{C}_i x_i x_i^T \hat{C}_i^T + B_i^T P_i x x^T \hat{C}_i^T = 0, \tag{4.5e}$$

hold for all $i = 1, \cdots, N$.

Proof. Substituting structured strategies (4.4) into performance indices (2.5) yields

$$J_i = \frac{1}{2} \|y(t_f)\|_{F_i}^2 + \frac{1}{2} \int_0^{t_f} [\|y\|_{Q_i}^2 + \|K_i^s \hat{C}_i x\|_{R_i}^2] dt. \tag{4.6}$$

The Hamiltonian is defined as

$$H_i = \frac{1}{2} \|x\|_{C^T Q_i C}^2 + \frac{1}{2} \|K_i^s \hat{C}_i x\|_{R_i}^2 + \lambda_i^T \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right) x,$$

where vector λ_i is the Lagrangian multiplier. According to the Pontryagin's minimum principle,

the necessary optimality conditions are

$$\dot{x} = \frac{\partial H_i}{\partial \lambda_i} = \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right) x, \tag{4.7a}$$

$$\dot{\lambda}_i = -\frac{\partial H_i}{\partial x} = -C^T Q_i Cx - \hat{C}_i^T (K_i^s)^T R_i K_i^s \hat{C}_i x - \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right)^T \lambda_i, \tag{4.7b}$$

$$\lambda_i(t_f) = C_i^T F_i C_i x(t_f), \tag{4.7c}$$

$$\frac{\partial H_i}{\partial K_i^s} = R_i K_i^s \hat{C}_i x x^T \hat{C}_i^T + B_i^T \lambda_i x^T \hat{C}_i^T = 0. \tag{4.7d}$$

for all $i = 1, \cdots, N$. Letting $\lambda_i = P_i x$ and substituting it into the above equations yields equations

(4.5). □

Obviously, if matrix $(\hat{C}_i x_i x_i^T \hat{C}_i^T)$ is invertible, then the optimal K can be obtained directly from condition (4.7d) as

$$K_i^s = -R_i^{-1} B_i^T \lambda_i x^T \hat{C}_i^T (\hat{C}_i x_i x_i^T \hat{C}_i^T)^{-1}. \tag{4.8}$$

Substituting (4.8) into (4.7a) and (4.7b) yields a highly nonlinear two-point boundary value problem, which is difficult to solve. Moreover, matrix $(C_i x_i x_i^T C_i^T)$ is generally not invertible. Therefore, in order to solve for $K_1^s, \cdots, K_N^s$, a gradient based iterative algorithm is proposed as follows.

Algorithm 4.1.

1. *Choose any time-varying feedback gains $K_1^{s0}(t), \cdots, K_N^{s0}(t)$ for $t \in [0, t_f]$ as the initial guessing.*

2. *At step k, substitute $K_i^s(t) = K_i^{sk}(t)$ for all $i = 1, \cdots, N$ into equations (4.5a) and (4.5c) and solve for $x(t)$ and $P_1(t), \cdots, P_N(t)$ for $t \in [0, t_f]$.*

3. If

$$\max_{0 \le t \le t_f} \|(R_i K_i^{sk} \hat{C}_i + B_i^T P_i) x x^T \hat{C}_i^T\|_2$$

is less than a stopping criteria, then $K_1^{sk}(t), \cdots, K_N^{sk}(t)$ are the solutions. Otherwise, go to
step 4.

4. Update $K_i^{sk}(t)$ according to

$$K_i^{s(k+1)}(t) = K_i^{sk}(t) - \epsilon_i \frac{\partial H_i}{\partial K_i^s}(K_i^{sk}) \quad \forall i = 1, \cdots, N, \tag{4.9}$$

where ϵ_i is the step size and $\frac{\partial H_i}{\partial K_i^s}(K_i^{sk})$ is defined in (4.7d) with K_i^s replaced with K_i^{sk}. Set
$k \to k + 1$ and go to step 2.

4.2.2   Noninferior Strategy Design

With parameterized $K_1^s, \cdots, K_N^s$ in (4.4), the noninferior solution is presented as the following
theorem based on the optimal output feedback approach.

Theorem 4.4. *For the differential game in an N-agent system under distributed feedback infor-*
mation structure defined by system dynamics (2.2) and performance indices (2.5), the strategies in

(4.4) form a feedback noninferior solution if the following equations holds:

$$\dot{x} = \bar{A}x, \tag{4.10a}$$

$$\bar{A} = \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right), \tag{4.10b}$$

$$\dot{P}_i + P_i \bar{A} + \bar{A}^T P_i + \sum_{j=1}^{N} \alpha_j [C^T Q_j C + \hat{C}_i^T (K_i^s)^T R_i K_i^s \hat{C}_i] = 0 \tag{4.10c}$$

$$P_i(t_f) = C_i^T F_i C \tag{4.10d}$$

$$\alpha_i R_i K_i^s \hat{C}_i x x^T \hat{C}_i^T + B_i^T \lambda x^T \hat{C}_i^T = 0, \tag{4.10e}$$

hold for all $i = 1, \cdots, N$.

Proof. Substituting structured strategies (4.4) into performance indices (2.5), the convex combination of the agents' performance indices becomes

$$J = \frac{1}{2} \sum_{j=1}^{N} \alpha_j \|y(t_f)\|_{F_j}^2 + \frac{1}{2} \int_0^{t_f} \sum_{j=1}^{N} \alpha_j [\|y\|_{Q_j}^2 + \|K_j^s \hat{C}_j x\|_{R_j}^2] \mathrm{d}t. \tag{4.11}$$

The Hamiltonian is defined as

$$H = \frac{1}{2} \sum_{j=1}^{N} \alpha_j (\|x\|_{C^T Q_j C}^2 + \|K_j^s \hat{C}_j x\|_{R_j}^2) + \lambda^T \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right) x,$$

where vector λ is the Lagrangian multiplier. According to the Pontryagin's minimum principle,

the necessary optimality conditions are

$$\dot{x} = \frac{\partial H}{\partial \lambda} = \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right) x, \tag{4.12a}$$

$$\dot{\lambda} = -\frac{\partial H}{\partial x} = -\sum_{j=1}^{N} \alpha_j [C^T Q_j C + \hat{C}_i^T (K_i^s)^T R_i K_i^s \hat{C}_i] x - \left(A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j \right)^T \lambda, \tag{4.12b}$$

$$\lambda(t_f) = \sum_{j=1}^{N} \alpha_j C_j^T F_j C_j x(t_f), \tag{4.12c}$$

$$\frac{\partial H}{\partial K_i^s} = \alpha_i R_i K_i^s \hat{C}_i x x^T \hat{C}_i^T + B_i^T \lambda x^T \hat{C}_i^T = 0 \tag{4.12d}$$

for all $i = 1, \cdots, N$. Letting $\lambda = Px$ and substituting it into the above equations yields equations (4.10). $\qquad\qquad\square$

Therefore, a gradient based iterative algorithm for deriving $K_1^s, \cdots, K_N^s$ that satisfy the conditions in (4.10) can be proposed in the same fashion as Algorithm 4.1.

There are several issues regarding the optimal output feedback based game strategy design approach, which are illustrated as follows:

1. It is important to point out that using the optimal output feedback design approach, the derived feedback gains $K_1^s, \cdots, K_N^s$ for both the Nash equilibrium and noninferior solution will depend upon the state, $x(t)$, as shown in equation (4.5e) and equation (4.10e). Moreover, it is equivalent to say that these feedback gains in fact depend upon the initial state, $x(0)$. Therefore, different sets of feedback matrices have to be derived given different initial states of the system. One way to overcome the initial state dependence is assuming that the initial state, x_0, is a random variable with certain probabilistic distribution [54]. Under this assumption, the agents will try to minimize the expected value of performance indices in (2.5). The initial state dependence is hence eliminated because the resulting feedback matri-

ces are only dependent on the covariance of the random initial state, which is assumed to be a given value.

2. For the differential game in a multi-agent system environment, since every agent's strategy has influence on other agents' performance indices values, the structured strategies defined in (4.4) with the feedback gains $K_1^s, \cdots, K_N^s$ need to be simultaneously parameterized and optimized with respect to the performance indices in order to obtain the distributed Nash equilibrium and noninferior solution. In other words, the optimal output feedback based approach requires all the agents to be able to choose the feedback gains freely. However, if there exists a certain constraint for the agent on its choice of the feedback matrix, then there is no way to apply the game strategy to each and every agent derived using optimal output feedback approach.

Realizing the existing issues in the optimal output feedback based approach, we propose a novel design approach for both Nash equilibrium and noninferior solution under distributed feedback information in the following section.

4.3   Best Achievable Performance Indices Approach

In this section, we present the distributed game strategy design based on a novel concept of best achievable performance indices. As shown in Section 4.1, with the structure constraints imposed on the feedback gain matrices, it is generally not possible for the agent's strategies form a Nash equilibrium or noninferior solution with respect to the original performance indices using the Riccati equation approach. However, in real life applications, the performance indices as a design criteria can usually be adjusted according to the real operation condition or situation. This inspires us to consider the game strategy design in a reverse manner. The basic idea of the approach

proposed in this section is as follows: For any structured strategies of the agents that conform to the information graph constraint, based on the inverse optimality, there exist a set of performance indices for the agents such that their structured strategies form a Nash equilibrium or noninferior solution. Among all the possible sets of performance indices, we find one set that is closest to the original set of indices. Therefore, the designed Nash equilibrium or noninferior solution is chosen to be the one corresponding to the closest set of performance indices. In what follows, we present the Nash strategy and noninferior strategy design approach based on the proposed concept of the best achievable performance index.

4.3.1 Nash Strategy Design

First of all, all the possible sets of performance indices with respect to which the structured strategies $u_1^s, \cdots, u_N^s$ described in (4.4) form a Nash equilibrium are presented as the following theorem based on inverse optimality.

Theorem 4.5. *For the differential game in an N-agent system under distributed feedback information structure defined by system dynamics (2.2) and performance indices (2.5), the strategies (4.4) form a Nash equilibrium with respect to the following performance indices:*

$$J_i^s = \frac{1}{2}\|x(t_f)\|_{C_i^T F_i C_i}^2 + \frac{1}{2}\int_0^{t_f}(\|x\|_{Q_i^s}^2 + x^T\Gamma_i^T u_i + u_i^T\Gamma_i x + \|u_i\|_{R_i}^2)dt \quad \forall i = 1, \cdots, N \quad (4.13)$$

where

$$Q_i^s = -A^T P_i - \dot{P}_i - P_i A + \hat{C}_i^T(K_i^s)^T R_i K_i^s \hat{C}_i + \sum_{j=1,j\neq i}^{N}(P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T(K_j^s)^T B_j^T P_i),$$
$$(4.14a)$$

$$\Gamma_i(K_i) = -R_i K_i \hat{C}_i - B_i^T P_i, \qquad (4.14b)$$

and P_i is any symmetric differentiable matrix with terminal condition $P_i(t_f) = C_i^T(t_f)F_iC_i(t_f)$.

Proof. Consider the Lyapunov function $V_i = 1/2x^T P_i x$ for agent i where matrix P_i is symmetric, differentiable, and satisfies the terminal condition $P_i(t_f) = C_i^T(t_f)F_iC_i(t_f)$. The derivative of V_i along the trajectory of system (2.2) is

$$\dot{V}_i = \frac{1}{2}(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + \frac{1}{2}x^T \dot{P}_i x + \frac{1}{2}x^T P_i(Ax + \sum_{j=1}^{N} B_j u_j). \tag{4.15}$$

Integrating the above equation from 0 to t_f yields

$$
\begin{aligned}
V_i(t_f) =& V_i(0) + \frac{1}{2}\int_0^{t_f} [(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + x^T \dot{P}_i x + x^T P_i(Ax + \sum_{j=1}^{N} B_j u_j)]\mathrm{d}t \\
=& V_i(0) + \frac{1}{2}\int_0^{t_f} [(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + x^T \dot{P}_i x + x^T P_i(Ax + \sum_{j=1}^{N} B_j u_j) \\
& + \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 - \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 + \sum_{j=1,j\neq i}^{N} x^T(P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T(K_j^s)^T B_j^T P_i)x \\
& - \sum_{j=1,j\neq i}^{N} x^T(P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T(K_j^s)^T B_j^T P_i)x]\mathrm{d}t, \\
=& V_i(0) + \frac{1}{2}\int_0^{t_f} \{-\|x\|_{Q_i^s}^2 - x^T \Gamma_i^T u_i - u_i^T \Gamma_i x - \|u_i\|_{R_i}^2 + \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 \\
& + \sum_{j=1,j\neq i}^{N} [(u_j - K_j^s \hat{C}_j x)^T B_j^T P_i x + x^T P_i B_j(u_j - K_j^s \hat{C}_j x)]\}\mathrm{d}t,
\end{aligned}
$$

where matrices Q_i^s and Γ_i are defined in (4.14). Hence, recalling the definition of the performance index in (4.13) and $P_i(t_f) = C_i^T(t_f)F_iC_i(t_f)$, one can obtain

$$
\begin{aligned}
J_i^s =& V_i(0) + \frac{1}{2}\int_0^{t_f} \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 \\
& + \sum_{j=1,j\neq i}^{N} [(u_j - K_j^s \hat{C}_j x)^T B_j^T P_i x + x^T P_i B_j(u_j - K_j^s \hat{C}_j x)]\mathrm{d}t. \tag{4.16}
\end{aligned}
$$

Since matrix R_i is positive definite, it is clear that performance index J_i^s in (4.16) reaches its minimum when $u_i^s = K_i^s \hat{C}_i x$ and $u_j^s = K_j^s \hat{C}_j x$ for $j \neq i$. Since the above analysis holds for all $i = 1, \cdots, N$, the inequality in (2.6) with J_i replaced by J_i^s holds if $u_i^* = K_i \hat{C}_i x$ for all $i = 1, \cdots, N$. Therefore, structured strategies (4.4) form a Nash equilibrium. □

Theorem 4.5 provides all the possible sets of performance indices (parameterized by feedback gain $K_1^s, \cdots, K_N^s$ and $P_1, \cdots, P_N$) with respect to which the structured strategies $u_1^s, \cdots, u_N^s$ expressed in (4.4) form a Nash equilibrium. For the convenience of the following analysis, we assume that matrices $P_1, \cdots, P_N$ are chosen to be the solutions to the coupled differential Riccati equations in (4.2). Comparing the set of performance indices in (4.13) with the set of original performance indices in (2.5), the differences between them are the values of matrices Q_i^s and Γ_i for all $i = 1, \cdots, N$. If $Q_i^s = C_i^T Q_i C_i$ and $\Gamma = 0$, then J_i^s becomes identical to J_i and structured strategy $u_1^s, \cdots, u_N^s$ form a Nash equilibrium with respect to the original performance indices. However, it is generally not possible to find proper values of $K_1^s, \cdots, K_N^s$ that achieve this. Therefore, one way is to find a set of performance indices among all possible performance indices in (4.13) which is closest to the ones in (2.5), that is, to make Q_i^s as close to $(C_i^T Q_i C_i)$ as possible and make Γ as close to 0 as possible. We call such set of performance index (closest to the original indices) the **best achievable performance indices**, which is defined formally as follows.

Definition 4.1. *Given the set of performance indices in (4.13) and the set of original performance indices in (2.5), if*

$$\int_0^{t_f} \|Q_i^s - C_i^T Q_i C_i\|_f^2 dt \quad and \quad \int_0^{t_f} \|\Gamma_i\|_f^2 dt \quad \forall i = 1, \cdots, N \qquad (4.17)$$

are simultaneously minimized by feedback matrices $K_i^s(t) = K_i^{s*}(t)$ for all $i = 1, \cdots, N$ where $\| \cdot \|_f$ is the Frobenius norm, then the set of performance index $J_1^{s*}, \cdots, J_N^{s*}$ corresponding to $K_1^{s*}(t), \cdots, K_N^{s*}(t)$ among all the sets of performance indices in (4.13) are called the best achiev-

able performance indices.

The concept of best achievable performance indices can be interpreted as a set of performance indices that is in the class of performance indices described by (4.13) while is also closest to the original indices (2.5) in terms of the Frobenius norm of the difference between the performance index coefficient matrices. Note that if matrix C_i is invertible, then both $\|Q_i^s - C_i^T Q_i C_i\|_f^2$ and $\|\Gamma_i\|_f^2$ can achieve the minimum value, zero, under the feedback matrix $K_i^{s*} = -R_i^{-1} B^T P_i \hat{C}_i^{-1}$. Substituting this feedback matrix into (4.14a) yields the differential Riccati equations (4.2). As such, the result reduces to the Nash equilibrium of the linear quadratic differential game. However, if matrix $\hat{C}$ is not invertible, finding $K_i^s = K_i^{s*}$ for all $i = 1, \cdots, N$ to minimize the terms in (4.17) simultaneously is quite difficult. Therefore, in order to find matrices $K_1^{s*}(t), \cdots, K_N^{s*}(t)$ corresponding to the best achievable performance indices, we need to solve a multi-objective optimization problem of minimizing $\int_0^{t_f} \|Q_i^s - C_i^T Q_i C_i\|_f^2 \mathrm{d}t$ and $\int_0^{t_f} \|\Gamma_i\|_f^2 \mathrm{d}t$ for all $i = 1, \cdots, N$ simultaneously. One way to accomplish this is to minimize a convex combination of these terms as follows.

$$\phi(K_1^s, \cdots, K_N^s) = \int_0^{t_f} H \mathrm{d}t, \tag{4.18}$$

where

$$H = \sum_{j=1}^{N} (\beta_{j1} \|Q_j^s - C_j^T Q_j C_j\|_f^2 + \beta_{j2} \|\Gamma_j\|_f^2) \tag{4.19}$$

where $0 \leq \beta_{j1} \leq 1$, $0 \leq \beta_{j2} \leq 1$, and $\sum_{j=1}^{N} (\beta_{j1} + \beta_{j2}) = 1$. The minimization problem reduces to finding matrices $K_1^{s*}(t), \cdots, K_N^{s*}(t)$ such that

$$\phi(K_1^{s*}(t), \cdots, K_N^{s*}(t)) \leq \phi(K_1^s(t), \cdots, K_N^s(t)) \quad \forall K_1^s(t), \cdots, K_N^s(t). \tag{4.20}$$

This minimization problem is generally quite difficult to solve analytically. A possible numerical

approach is using gradient based iterative algorithms [74]. Since these algorithms will require an expression for the gradient of $H(t)$ with respect to $K_1^s(t), \cdots, K_N^s(t)$. Recalling the property of the Frobenius norm $\|S\|_f^2 = \text{Tr}(S^T S)$ where $\text{Tr}(\cdot)$ is the matrix trace operation, equation (4.19) becomes

$$H_i = \sum_{j=1}^{N} \{\beta_{j1}\text{Tr}[(Q_j^s - C_j^T Q_j C_j)^2] + \beta_{j2}\}\text{Tr}(\Gamma_j \Gamma_j^T).$$

Hence, the partial derivatives of H with respect to $K_1^s, \cdots, K_N^s$ are

$$\nabla_{K_i^s} H = 4\beta_{i1} R_i K_i^s \hat{C}_i (Q_i^s - C_i^T Q_i C_i)\hat{C}_i^T$$
$$+ 4 \sum_{j=1, j\neq i}^{N} \beta_{j1}[B_i^T P_j (Q_j^s - C_j^T Q_j C_j)\hat{C}_i^T] + 2\beta_{j2} R_i \Gamma_i \hat{C}_i^T \quad \forall i = 1, \cdots, N. \quad (4.21)$$

The following gradient based iterative algorithm is proposed:

Algorithm 4.2.

1. *Choose $K_1^{s0}(t), \cdots, K_N^{s0}(t)$ for $t \in [0, t_f]$ as the initial guessing.*

2. *If*

$$\max_{0 \leq t \leq t_f, i=1, \cdots, N} \|\nabla_{K_i^s} H(K_1^{sk}, \cdots, K_N^{sk})\|_2$$

   is less than a stopping criteria where $\nabla_{K_i^s} H(K_1^{sk}, \cdots, K_N^{sk})$ is defined in (4.21) with $K_i^s(t) = K_i^{sk}(t)$ for all $i = 1, \cdots, N$, then $K_1^{sk}(t), \cdots, K_N^{sk}(t)$ are the solutions. Otherwise, go to step 3.

3. *Update $K_1^{sk}(t), \cdots, K_N^{sk}(t)$ according to*

$$K_i^{s(k+1)}(t) = K_i^{sk}(t) - \epsilon_i \nabla_{K_i^s} H(K_1^{sk}, \cdots, K_N^{sk}) \quad \forall i = 1, \cdots, N$$

63

where ϵ_i is the step size. Set $k \to k+1$ and go to step 2.

Note that by varying the coefficient $\beta_{11}, \cdots, \beta_{N1}$ and $\beta_{12}, \cdots, \beta_{N2}$, a noninferior set of the solutions can be generated. An appropriate choice of these coefficients can be made to place a desired emphasis on minimizing each and every term in (4.17).

4.3.2 Noninferior Strategy Design

The same idea of the Nash strategy design can be also applied to the noninferior strategy design. First of all, all the possible sets of performance indices with respect to which the structured strategies $u_1^s, \cdots, u_N^s$ described in (4.4) form a noninferior solution are presented as the following theorem based on inverse optimality.

Theorem 4.6. *For the differential game in an N-agent system under distributed feedback information structure defined by system dynamics (2.2) and performance indices (2.5), the strategies in (4.4) form a noninferior solution with respect to the following performance indices*

$$J_i^s = \frac{1}{2}\|x(t_f)\|_{C_i^T F_i C_i}^2 + \frac{1}{2}\int_0^{t_f}(\|x\|_{Q_i^s}^2 + x^T\Gamma_i^T u_i + u_i^T\Gamma_i x + \|u_i\|_{R_i}^2)dt \quad \forall i = 1, \cdots, N \quad (4.22)$$

where

$$Q_i^s = -A^T P - \dot{P} - PA + \hat{C}_i^T(K_i^s)^T R_i K_i^s \hat{C}_i, \tag{4.23a}$$

$$\Gamma_i(K_i^s) = -R_i K_i^s \hat{C}_i - \frac{1}{\alpha_i} B_i^T P, \tag{4.23b}$$

and P is any symmetric differentiable matrix with $P(t_f) = \sum_{j=1}^N \alpha_j C_j^T(t_f) F_j C_j(t_f)$ and $0 \le \alpha_i \le 1$ for all $i = 1, \cdots, N$ and $\sum_{j=1}^N \alpha_j = 1$.

Proof. Consider the Lyapunov function $V = 1/2x^T Px$. The derivative of V along the trajectory of system (2.2) is the same as (4.15). Integrating (4.15) from 0 to t_f yields

$$V(t_f) = V(0) + \frac{1}{2}\int_0^{t_f}[(Ax + \sum_{j=1}^N B_j u_j)^T Px + x^T \dot{P}x$$
$$+ x^T P(Ax + \sum_{j=1}^N B_j u_j)]\mathrm{d}t$$

$$\frac{1}{2}x^T(t_f)\left(\sum_{j=1}^N \alpha_j C_j^T F_j C_j\right)x(t_f) = V(0) + \frac{1}{2}\int_0^{t_f}[(Ax + \sum_{j=1}^N B_j u_j)^T Px + x^T \dot{P}x$$
$$+ x^T P(Ax + \sum_{j=1}^N B_j u_j)$$
$$+ \sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2 - \sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2]\mathrm{d}t,$$
$$= V(0) + \frac{1}{2}\int_0^{t_f}\sum_{j=1}^N \alpha_j[-\|x\|_{Q_j^s}^2 - x^T \Gamma_j^T u_j - u_j^T \Gamma_j x$$
$$- \|u_j\|_{R_j}^2 + \|u_i - K_i^s \hat{C}_i x_i\|_{R_j}^2]\mathrm{d}t,$$

where matrices Q_i^s and Γ_i are defined in (4.23). Hence, denoting that $J^s = \sum_{j=1}^N J_j^s$ where J_j^s is defined in (4.22), we have

$$J^s = V(0) + \frac{1}{2}\int_0^{t_f}\sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2 \mathrm{d}t.$$

Since $\alpha_j \geq 0$ and matrix R_j is positive definite for all $j = 1, \cdots, N$, it is clear from the above equation that $u_j^s = K_j^s \hat{C}_j x$ for all $j = 1, \cdots, N$ are the optimal strategies and hence form a noninferior solution for any given $\alpha_1, \cdots, \alpha_N$. $\square$

Note that only if the parameters $\alpha_1, \cdots, \alpha_N$ are given, the expression of performance indices $J_1^s, \cdots, J_N^s$ in (4.22) can be obtained for any structured strategies $u_1^s, \cdots, u_N^s$ in (4.4). For the

convenience of the following analysis, we assume that the value of matrix P are chosen to be the solution to the differential Riccati equation in (3.5). The definition of the best performance indices in this case is in fact the same as Definition 4.1 with matrices Q_i^s and Γ_i^s replaced by the ones in (4.23). Similar to the Nash strategy design, to find matrices $K_1^{s*}, \cdots, K_N^{s*}$ corresponding to the best achievable performance indices, we solve a minimization problem with respect to the convex combination in (4.18). In this case, the partial derivatives of H with respect to $K_1^s, \cdots, K_N^s$ become

$$\nabla_{K_i^s} H = 4\beta_{i1} R_i K_i^s \hat{C}_i (Q_i^s - C_i^T Q_i C_i) \hat{C}_i^T + 2\beta_{j2} R_i \Gamma_i \hat{C}_i^T \quad \forall i = 1, \cdots, N \qquad (4.24)$$

where matrices Q_i^s and Γ_i are defined in (4.23). Therefore, with the expression of the gradient in (4.24), the same gradient based iterative algorithm as Algorithm 4.2 can be utilized to derive $K_1^{s*}, \cdots, K_N^{s*}$ corresponding to the best achievable performance indices algorithm for any given set of parameters $\alpha_1, \cdots, \alpha_N$. Again, by varying the coefficient $\beta_{11}, \cdots, \beta_{N1}$ and $\beta_{12}, \cdots, \beta_{N2}$, a noninferior set of the solutions can be generated. An appropriate choice of these coefficients can be made to place a desired emphasis on minimizing each and every term in (4.17).

With the best achievable performance indices approach proposed, in what follows, we point out several features of this approach and how this approach overcomes the issues in the previous optimal output feedback based approach.

1. Matrices $K_1^{s*}, \cdots, K_N^{s*}$ derived using the best achievable performance indices approach for both the Nash equilibrium and noninferior solution are independent on the initial state.

2. The optimal output feedback approach requires every agent to be able to choose the feedback gain matrices freely and optimize the parameterized feedback gains $K_1^s, \cdots, K_N^s$ simultaneously. While using the best performance indices approach, no such requirement is needed.

Specifically, if the structured strategies for one or more agents are fixed, then the best achievable performance indices can still be applicable in the sense that the rest agents can optimize their feedback matrices such that the set of performance indices with respect to which all the agents' strategies (the optimized strategies and the fixed strategies) form a Nash equilibrium or noninferior solution are closest to the original set of performance indices.

4.4 Feedback Game Strategy Design over the Infinite Time Horizon

In this section, we consider the feedback Nash strategy and noninferior strategy design for the differential games over the infinite time horizon in the multi-agent systems under distributed feedback information structure.

For the game over the infinite time horizon, we assume that dynamics (2.1) for every agent is time-invariant, stablizable, and detectable. The performance indices are given by

$$J_i = \frac{1}{2} \int_0^\infty (\|y\|_{Q_i}^2 + \|u_i\|_{R_i}^2) \mathrm{d}t \quad \forall i = 1, \cdots, N \tag{4.25}$$

where matrix Q_i is time-invariant and positive semi-definite (with exceptions in cases such as the pursuit-evasion games) and matrix R_i is time-variant and positive definite for all $i = 1, \cdots, N$. If the strategies of all the agents are constrained to be (4.4), then the following result similar result to Theorem 4.5 is obtained.

Theorem 4.7. *For the differential game in an N-agent system under distributed feedback information structure defined by system dynamics (2.2) and performance indices (4.25), the strategies in (4.4) with form a Nash equilibrium with respect to the following performance indices:*

$$J_i^s = \frac{1}{2} \int_0^\infty (\|x\|_{Q_i^s}^2 + x^T \Gamma_i^T u_i + u_i^T \Gamma_i x + \|u_i\|_{R_i}^2) dt \quad \forall i = 1, \cdots, N \tag{4.26}$$

67

if the closed-loop system $\dot{x} = (A + \sum_{j=1}^{N} B_j K_j^s \hat{C}_j)x$ is asymptotically stabilize, where

$$Q_i^s = -A^T P_i - P_i A + \hat{C}_i^T (K_i^s)^T R_i K_i \hat{C}_i + \sum_{j=1,j\neq i}^{N} (P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T (K_j^s)^T B_j^T P_i), \quad (4.27a)$$

$$\Gamma_i = -R_i K_i^s \hat{C}_i - B_i^T P_i, \quad (4.27b)$$

and matrix P_i is symmetric.

Proof. Consider the Lyapunov function $V_i = 1/2 x^T P_i x$ for agent i. Its derivative along the trajectory of (2.1) is

$$\dot{V}_i = \frac{1}{2}(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + \frac{1}{2}x^T P_i(Ax_i + \sum_{j=1}^{N} B_j u_j). \quad (4.28)$$

If the closed-loop system is asymptotically stable, integrating (4.28) from 0 to ∞ yields

$$\begin{aligned}
V_i(\infty) = 0 = &V_i(0) + \frac{1}{2}\int_0^\infty [(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + x^T P_i(Ax + \sum_{j=1}^{N} B_j u_j)]\mathrm{d}t \\
= &V_i(0) + \frac{1}{2}\int_0^\infty [(Ax + \sum_{j=1}^{N} B_j u_j)^T P_i x + x^T P_i(Ax + \sum_{j=1}^{N} B_j u_j) \\
&+ \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 - \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 + \sum_{j=1,j\neq i}^{N} x^T (P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T (K_j^s)^T B_j^T P_i)x \\
&- \sum_{j=1,j\neq i}^{N} x^T (P_i B_j K_j^s \hat{C}_j + \hat{C}_j^T (K_j^s)^T B_j^T P_i)x]\mathrm{d}t, \\
= &V_i(0) + \frac{1}{2}\int_0^\infty \{-\|x\|_{Q_i^s}^2 - x^T \Gamma_i^T u_i - u_i^T \Gamma_i x - \|u_i\|_{R_i}^2 + \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 \\
&+ \sum_{j=1,j\neq i}^{N} [(u_j - K_j^s \hat{C}_j x)^T B_j^T P_i x + x^T P_i B_j(u_j - K_j^s \hat{C}_j x)]\}\mathrm{d}t,
\end{aligned}$$

where matrices Q_i^s and Γ_i are defined in (4.27). Hence, recalling the definition of the performance

index in (4.26), we have

$$
\begin{aligned}
J_i^s =& V_i(0) + \frac{1}{2} \int_0^\infty \|u_i - K_i^s \hat{C}_i x\|_{R_i}^2 \\
&+ \sum_{j=1, j \neq i}^N [(u_j - K_j^s \hat{C}_j x)^T B_j^T P_i x + x^T P_i B_j (u_j - K_j^s \hat{C}_j x)] \mathrm{d}t..
\end{aligned}
\tag{4.29}
$$

Since matrix R_i is positive definite, the inequalities in (2.6) hold true with J_i replaced by J_i^s if $u_i^* = K_i^s \hat{C}_i x$ for $i = 1, \cdots, N$. Therefore, the structured strategies in (4.4) form a Nash equilibrium. $\qquad\square$

For the convenience of the following derivation, the values of matrices $P_1, \cdots, P_N$ are chosen to be the solutions to the algebraic version of the coupled differential Riccati equations in (4.2) by setting $\dot{P}_1 = \cdots = \dot{P}_N = 0$. Therefore, with all the possible performance indices parameterized in (4.26), the definition of best achievable performance indices over the infinite time horizon is as follows:

Definition 4.2. *Given the set of performance indices in (4.26) and the set of original performance indices in (4.25) over the infinite time horizon, if*

$$
\|Q_i^s - C_i^T Q_i C_i\|_f^2 \quad \text{and} \quad \|\Gamma_i\|_f^2 \quad \forall i = 1, \cdots, N
\tag{4.30}
$$

are simultaneously minimized by feedback matrices $K_i^s(t) = K_i^{s*}(t)$ for all $i = 1, \cdots, N$ where $\|\cdot\|_f$ is the Frobenius norm, then the set of performance index $J_1^{s*}, \cdots, J_N^{s*}$ corresponding to $K_1^{s*}(t), \cdots, K_N^{s*}(t)$ among all the sets of performance indices in (4.26) are called the best achievable performance indices.

Note that if matrix $\hat{C}_i$ is invertible, both $\|Q_i^s - C_i^T Q_i C_i\|_f^2$ and $\|\Gamma_i\|_f^2$ can achieve the minimum value, 0, under the feedback matrix $K_i^s = -R_i^{-1} B^T P_i \hat{C}_i^{-1}$. Substituting this feedback matrix

into (4.27a) yields the algebraic version of differential Riccati equations (4.2). In the case that matrix $\hat{C}_i$ is not invertible, we need to utilize a numerical algorithm to find feedback matrices $K_1^{s*}, \cdots, K_N^{s*}$ corresponding to the set of best achievable performance indices. Toward that end, we define a convex combination of the terms in (4.30) as an objective function to minimize, which is the same as H in (4.19). Hence, with the same expression of partial derivatives of H with respect to $K_1^s, \cdots, K_N^s$ in (4.21), the following gradient based iterative algorithm similar to Algorithm 4.2 is proposed to derive the feedback matrices $K_1^{s*}, \cdots, K_N^{s*}$.

Algorithm 4.3.

1. *Choose $K_1^{s0}(t), \cdots, K_N^{s0}(t)$ as the initial guessing such that the closed loop system matrix $(A + \sum_{j=1}^{N} B_j K_i^{s0} \hat{C}_i)$ is asymptotically stable.*

2. *If $(A_i + \sum_{j=1}^{N} B_j K_j^{sk} \hat{C}_j)$ is asymptotically stable and $\| \max_{i=1,\cdots,N} \nabla_{K_i^{sk}} H_i(K_1^{sk}, \cdots, K_N^{sk}) \|_f$ is less than a stopping criteria where $\nabla_{K_i^s} H(K_1^{sk}, \cdots, K_N^{sk})$ is defined in (4.21) with Q_i^s and Γ_i defined in (4.27) and $K_i^s(t) = K_i^{sk}(t)$ for all $i = 1, \cdots, N$, then $K_1^{sk}(t), \cdots, K_N^{sk}(t)$ are the solutions. Otherwise, go to step 3.*

3. *Update $K_1^{sk}(t), \cdots, K_N^{sk}(t)$ according to*

$$K_i^{s(k+1)}(t) = K_i^{sk}(t) - \epsilon_i \nabla_{K_i^s} H(K_1^{sk}, \cdots, K_N^{sk}) \quad \forall i = 1, \cdots, N$$

where ϵ_i is the step size. Set $k \to k + 1$ and go to step 2.

By varying the coefficients β_1, a noninferior set of the solutions can be generated and an appropriate choice can be made.

The same idea can be applied to the noninferior strategy design over the infinite time horizon. First of all, we have the following theorem similar to (4.8).

70

Theorem 4.8. *For the differential game in an N-agent system under distributed feedback infor-mation structure defined by system dynamics (2.2) and performance indices (2.5), the strategies in (4.4) form a noninferior solution with respect to the following performance indices*

$$J_i^s = \frac{1}{2} \int_0^\infty (\|x\|_{Q_i^s}^2 + x^T \Gamma_i^T u_i + u_i^T \Gamma_i x + \|u_i\|_{R_i}^2) dt \quad \forall i = 1, \cdots, N \tag{4.31}$$

if the closed-loop system $\dot{x} = (A + \sum_{j=1}^N B_j K_j^s \hat{C}_j) x$ is asymptotically stabilize, where

$$Q_i^s = -A^T P - PA + \hat{C}_i^T (K_i^s)^T R_i K_i^s \hat{C}_i, \tag{4.32a}$$

$$\Gamma_i(K_i^s) = -R_i K_i^s \hat{C}_i - \frac{1}{\alpha_i} B_i^T P, \tag{4.32b}$$

matrix P is symmetric, $0 \le \alpha_i \le 1$ for all $i = 1, \cdots, N$, and $\sum_{j=1}^N \alpha_j = 1$.

Proof. Consider the Lyapunov function $V = 1/2 x^T P x$. The derivative of V along the trajectory of system (2.2) is the same as (4.28). If the closed-loop system is asymptotically stable, integrating (4.28) from 0 to ∞ yields

$$
\begin{aligned}
V(\infty) = 0 =& V(0) + \frac{1}{2} \int_0^{t_f} [(Ax + \sum_{j=1}^N B_j u_j)^T P x + x^T \dot{P} x + x^T P (Ax + \sum_{j=1}^N B_j u_j)] \mathrm{d}t \\
=& V(0) + \frac{1}{2} \int_0^{t_f} [(Ax + \sum_{j=1}^N B_j u_j)^T P x + x^T \dot{P} x + x^T P (Ax + \sum_{j=1}^N B_j u_j) \\
& + \sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2 - \sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2] \mathrm{d}t, \\
=& V(0) + \frac{1}{2} \int_0^{t_f} \sum_{j=1}^N \alpha_j [-\|x\|_{Q_j^s}^2 - x^T \Gamma_j^T u_j - u_j^T \Gamma_j x \\
& - \|u_j\|_{R_j}^2 + \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2] \mathrm{d}t,
\end{aligned}
$$

where matrices Q_i^s and Γ_i are defined in (4.32). Hence, denoting that $J^s = \sum_{j=1}^N J_j^s$ where J_j^s is

71

defined in (4.31), we have

$$J^s = V(0) + \frac{1}{2} \int_0^{t_f} \sum_{j=1}^N \alpha_j \|u_j - K_j^s \hat{C}_j x\|_{R_j}^2 \, \mathrm{d}t.$$

Since $\alpha_j \geq 0$ and matrix R_j is positive definite for all $j = 1, \cdots, N$, it is clear from the above equation that $u_j^s = K_j^s \hat{C}_j x$ for all $j = 1, \cdots, N$ are the optimal strategies and hence form a noninferior solution for any given $\alpha_1, \cdots, \alpha_N$. $\qquad\square$

For convenience, the values of matrix P in (4.32) is chosen to be the solution to the algebraic version of the differential Riccati equation in (3.5) by setting $\dot{P} = 0$. The definition of best achievable performance indices in this case is the same as Definition 4.2 with matrices Q_i^s and Γ_i defined in (4.32). with the same expression of partial derivatives of H with respect to $K_1^s, \cdots, K_N^s$ in (4.24), the same gradient based iterative algorithm as Algorithm 4.3 can be utilized to derive $K_1^{s*}, \cdots, K_N^{s*}$ corresponding to the best achievable performance indices algorithm for any given set of parameters $\alpha_1, \cdots, \alpha_N$. Again, a noninferior set of the solutions can be generated by varying the coefficients β_1 and an appropriate choice can be made.

Note that in the Algorithm 4.3, the initial guessing has to be a stabilizing feedback gain and the stability of the closed-loop system is verified at every iteration to make sure that the closed-loop system under the resulting control is asymptotically stable. Therefore, we are confronted with two issues. One is that whether the system can be stabilized by the structured strategies in the form of (4.4), and the other is how to find an initial stabilizing control to start the algorithm if the system is stabilizable. One possible approach is to utilize the Lyapunov stability criterion, which is to find solution to the following Lyapunov inequality

$$P\left(A + \sum_{j=1}^N B_j K_j^s \hat{C}_j\right) + \left(A + \sum_{j=1}^N B_j K_j^s \hat{C}_j\right)^T P < 0, \qquad (4.33)$$

There exist many numerical approaches to solve this problem, one of which is based on the linear matrix inequality (LMI) technique. The inequality above can be converted into a LMI feasibility problem and there exist many software tools to solve the LMI feasibility problem efficiently such as YALMIP [75]. Moreover, if the LMIs are feasible, the software program will automatically generate a feasible solution, which can be used as the initial stabilizing feedback gain to initialize Algorithm 4.3. Please refer to [76] for more details for a possible approach.

In this chapter, we considered the game strategy design approach in the multi-agent system under distributed feedback information structure. The basic idea of this approach is to design structured feedback strategy for the agents such that these strategies form a Nash equilibrium or noninferior solution with respect to a set of performance indices that are closest to the original indices. This approach overcomes several shortcomings in the conventional optimal feedback based approach and is extended to the game over the infinite time horizon.

CHAPTER 5: APPLICATIONS

In this chapter, two application examples of the differential games in multi-agent systems are considered. One is the unmanned aerial vehicles (UAVs) formation control problem and the other is the multi-pursuer single-evader differential game with limited observations. The former one is solved using the proposed open-loop noninferior strategy design approach and the latter one is solved using the best achievable performance indices approach.

5.1　UAV Formation Control Using Differential Game Approach

In this section, we consider N UAVs that are trying to form a prescribed formation and design open-loop controls for each and every UAV to achieve this objective. The point-mass dynamics of UAVs are modeled as follows [77] and are shown in Figure 5.1:

$$\dot{x}_i = V_i \cos \gamma_i \cos \chi_i, \tag{5.1a}$$

$$\dot{y}_i = V_i \cos \gamma_i \sin \chi_i \tag{5.1b}$$

$$\dot{h}_i = V \sin \gamma_i \tag{5.1c}$$

$$\dot{V}_i = \frac{T_i - D_i}{m_i} - g \sin \gamma_i \tag{5.1d}$$

$$\dot{\gamma}_i = \frac{L \cos \phi_i - m_i g \cos \gamma_i}{m_i V_i} \tag{5.1e}$$

$$\dot{\chi}_i = \frac{L_i \sin \phi_i}{m_i V_i \cos \gamma_i} \tag{5.1f}$$

for $i = 1, \cdots, N$, where x_i is the down-range displacement, y_i is the cross-range displacement, h_i is the altitude, V_i is the ground speed which is assumed to be equal to the airspeed in this paper, γ_i is the flight path angle, χ_i is the heading angle, T_i is the engine thrust, D_i is the drag, m_i is the

UAV mass, g is the acceleration due to gravity, L_i is the lift, and ϕ_i is the banking angle. The three control inputs of UAV i is the banking angle ϕ_i, lift L_i, and engine thrust T_i.
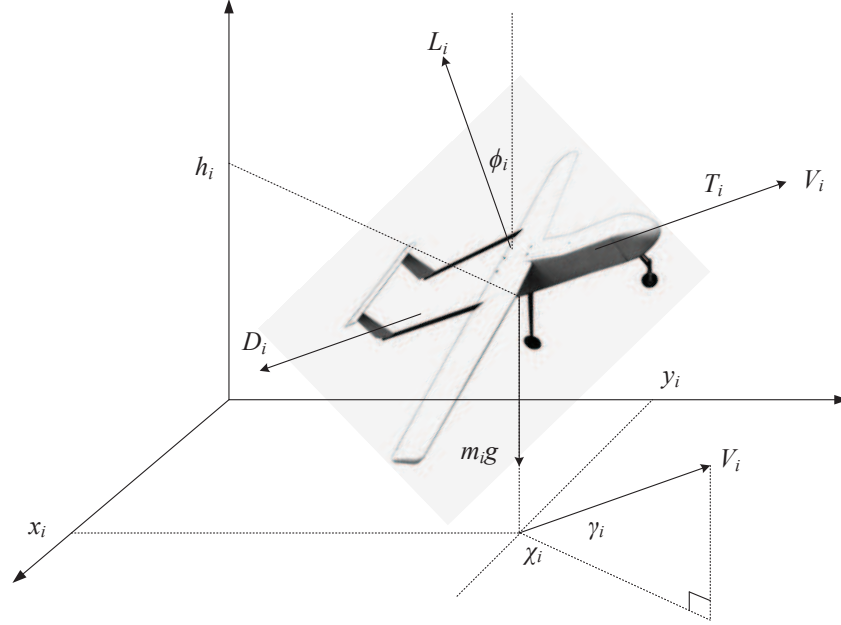

Figure 5.1: UAV Model

It is shown in [77] that the highly nonlinear UAV model in (5.1) can be pre-linearized using feedback linearization to be

$$\ddot{x}_i = u_{xi}, \quad \ddot{y}_i = u_{yi}, \quad \ddot{h}_i = u_{hi} \tag{5.2}$$

where u_{xi}, u_{yi}, and u_{hi} are the virtual acceleration control inputs. These virtual control inputs and

the real control inputs are related through the following equations

$$\phi_i = \tan^{-1}\left(\frac{u_{yi}\cos\chi_i - u_{xi}\sin\chi_i}{(u_{hi}+g)\cos\gamma_i - (u_{xi}\cos\chi_i + u_{yi}\sin\chi_i)\sin\gamma_i}\right) \tag{5.3a}$$

$$L_i = m_i\frac{(u_{hi}+g)\cos\gamma_i - (u_{xi}\cos\chi_i + u_{yi}\sin\chi_i)\sin\gamma_i}{\cos\phi_i} \tag{5.3b}$$

$$T_i = m_i[(u_{hi}+g)\sin\gamma_i + (u_{xi}\cos\chi_i + u_{yi}\sin\chi_i)\cos\gamma_i] + D_i \tag{5.3c}$$

where $\tan\chi_i = \dot{y}_i/\dot{x}_i$ and $\sin\gamma_i = \dot{h}_i/V_i$. Therefore, after the virtual control inputs are designed based on the linear model (5.2), the real control inputs can then be obtained by substituting the virtual ones into equations in (5.3). Expressing (5.2) in terms of state-space representation yields

$$\dot{z}_i = Az_i + Bu_i, \tag{5.4a}$$

$$p_i = C_p z_i \tag{5.4b}$$

$$v_i = C_v z_i \tag{5.4c}$$

where $z_i = [p_i^T \quad v_i]^T$ is the state vector, p_i is the position vector, v_i is the velocity vector, $u_i = [u_{x_i}^T \ u_{y_i}^T \ u_{h_i}^T]^T$ is the virtual acceleration control vector,

$$A_i = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes I_3, \quad B_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes I_3, \quad C_p = [1\ 0] \otimes I_3, \quad C_v = [0\ 1] \otimes I_3,$$

Suppose that individual UAVs are able to communicate with each other according to a directed information graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. To achieve the formation requirement, graph $\mathcal{G}$ is assumed to be connected. Since the objective of the UAVs is to form a prescribed formation, assuming that the desired displacement vector pointing from UAV j to UAV i is μ_{ij}, the formation requirement can

be expressed mathematically in terms of the following performance index for UAV i to minimize:

$$J_i = \sum_{e_{ij} \in \mathcal{E}} \frac{1}{2}[\|p_i(t_f) - p_j(t_f) - \mu_{ij}\|^2 + \|v_i(t_f) - v_j(t_f)\|^2] + \frac{r_i}{2} \int_0^{t_f} \|u_i\|^2 \mathrm{d}t \qquad (5.5)$$

for all $i = 1, \cdots, N$, where $\| \cdot \|$ is the Euclidean norm or distance and r_i is a positive scalar. Performance index (5.5) means that UAV i will try to minimize the sum of the terminal formation errors and terminal velocity errors according to the information graph while at the same time minimizing its control effort made during the entire process. The larger r_i is, the larger penalty is placed on the control effort. Note that coefficients $r_1, \cdots, r_N$ are not necessarily the same because the choices of these coefficients reflect the real situation. For instance, if UAV i has sufficient fuel in its tank, it will naturally choose a small value of r_i in order to keep the desired formation with others actively. On the contrary, if UAV i does not have much fuel left in its tank, it will naturally choose a large value of r_i to preserve its energy or fuel cost. Therefore, we assume that the UAVs will play a cooperative game and collaborate with each other as a team to achieve the prescribed formation. This leads to solving the multi-objective optimization problem in (3.3) and finding the noninferior solution of the game. Toward that end, similar to (3.9), we define new state vectors as

$$s_{pi}(t) = [1 \ \ (t_f - t)]z_i(t) \quad \text{and} \quad s_{vi}(t) = [0 \ \ 1]z_i(t). \qquad (5.6)$$

Differentiating both sides of (5.6) with respect to t and recalling system dynamics (2.2) yield

$$\dot{s}_{pi} = \tilde{B}_p u_i \quad \text{and} \quad \dot{s}_{vi} = \tilde{B}_v u_i, \qquad (5.7)$$

where $\tilde{B}_p = (t_f - t)I_3$ and $\tilde{B}_v = I_3$. Based on the properties $s_{pi}(t_f) = p_i(t_f)$ and $s_{vi}(t_f) = v_i(t_f)$

for s_{pi} and s_{vi} defined in (5.6), the performance indices in (5.5) can be rewritten as

$$J_i = \sum_{e_{ij} \in \mathcal{E}} \frac{1}{2}[\|s_{pi}(t_f) - s_{pj}(t_f) - \mu_{ij}\|^2 + \|s_{vi}(t_f) - s_{vj}(t_f)\|^2] + \frac{r_i}{2} \int_0^{t_f} \|u_i\|^2 \mathrm{d}t. \qquad (5.8)$$

for all $i = 1, \cdots, N$. The convex combination in (3.3) with J_i defined in (5.8) can be expressed as

$$\begin{aligned} J = &\sum_{j=1}^{N} \frac{\alpha_j}{2} \sum_{e_{jk} \in \mathcal{E}} [\|s_{pj}(t_f) - s_{pk}(t_f) - \mu_{jk}\|^2 + \|s_{vj}(t_f) - s_{vk}(t_f)\|^2] \\ &+ \sum_{j=1}^{N} \frac{\alpha_j r_j}{2} \int_0^{t_f} \|u_j\|^2 \mathrm{d}t \quad \forall i = 1, \cdots, N, \end{aligned} \qquad (5.9)$$

where $\alpha_j > 0$ for all $j = 1, \cdots, N$. Similarly to (2.4), we define the Laplacian matrix $\mathcal{L} = [\mathcal{L}_{ij}] \in \mathbb{R}^{N \times N}$ associated with the graph among the N UAVs as follows:

$$\mathcal{L}_{ij} = \begin{cases} -(\alpha_i + \alpha_j) & \text{if } e_{ij} \in \mathcal{E} \text{ for } j \neq i \\ 0 & \text{if } e_{ij} \notin \mathcal{E} \text{ for } j \neq i \\ -\sum_{q=1, q \neq i}^{N} \mathcal{L}_{iq} & \text{if } j = i \end{cases}, \qquad (5.10)$$

It is obvious that $\mathcal{L}^T = \mathcal{L}$ because the graph is assumed to be undirected and hence matrix $\mathcal{L}$ is positive semi-definite. Before we present the result of open-loop noninferior solution, the following lemma is introduced.

Lemma 5.1. *All the eigenvalues of matrix M defined by*

$$M = [I_{2N} + W \otimes (R^{-1}\mathcal{L})] \otimes I_3, \qquad (5.11)$$

has positive real parts, where

$$W = \begin{bmatrix} w_{pp} & w_{pv} \\ w_{vp} & w_{vv} \end{bmatrix}, \tag{5.12a}$$

$$w_{pp} = \int_0^{t_f} \tilde{B}_p \tilde{B}_p^T dt = \frac{t_f^3}{3}, \quad w_{pv} = \int_0^{t_f} \tilde{B}_p \tilde{B}_v^T dt = \frac{t_f^2}{2} \tag{5.12b}$$

$$w_{vp} = \int_0^{t_f} \tilde{B}_v \tilde{B}_p^T dt = \frac{t_f^2}{2}, \quad w_{vv} = \int_0^{t_f} \tilde{B}_v \tilde{B}_v^T dt = t_f, \tag{5.12c}$$

$$R = diag\{\mu_1 r_1, \cdots, \mu_N r_N\}, \tag{5.12d}$$

and diag$\{\cdot\}$ stands for "diagonal matrix".

Proof. Firstly, since it is obvious that matrix W defined by (5.12a)-(5.12c) is positive define, all its eigenvalues are positive. Secondly, since matrix R in (5.12d) is a positive diagonal matrix, the product of $(R^{-1}\mathcal{L})$ becomes a new weighted Laplacian matrix whose eigenvalues still have non-negative real parts. Thirdly, since the eigenvalues of matrices' Kronecker product are the product of these matrices' eigenvalues, all the eigenvalues of matrix $[W \otimes (R^{-1}\mathcal{L}) \otimes I_3]$ have non-negative real parts. Therefore, all the eigenvalues of M in (3.14a) have positive real parts. $\square$

The open-loop Nash equilibrium solution is now presented as the following theorem similar to Theorem 3.3.

Theorem 5.1. *Given the differential game among N UAVs with system dynamics (5.7) and performance indices (5.8), the strategies*

$$u_i^* = -\frac{1}{\alpha_i r_i} F_i M^{-1} \left(\begin{bmatrix} s_p(0) \\ s_v(0) \end{bmatrix} + W_\mu \mu \right) + \frac{1}{\alpha_i r_i} \tilde{B}_p^T \mu_i \quad \forall i = 1, \cdots, N \tag{5.13}$$

79

form an open-loop Nash equilibrium, where matrix M is defined in (3.14a),

$$s_p = [s_{p1}^T \quad \cdots \quad s_{pN}^T]^T, \quad s_v = [s_{v1}^T \quad \cdots \quad s_{vN}^T]^T, \tag{5.14a}$$

$$F_i = [\tilde{B}_p^T \quad \tilde{B}_v^T][I_2 \otimes (d_i^T \mathcal{L}) \otimes I_3], \tag{5.14b}$$

$$W_\mu = \begin{bmatrix} w_{pp} \\ w_{vp} \end{bmatrix} \otimes R^{-1} \otimes I_3, \tag{5.14c}$$

$$\mu = \begin{bmatrix} \mu_1^T \quad \cdots \quad , \mu_N^T \end{bmatrix}^T, \tag{5.14d}$$

$$\mu_i = \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)\mu_{ij} \quad \forall i = 1, \cdots, N, \tag{5.14e}$$

$\mathcal{L}$ is the Laplacian matrix defined in (5.10), $d_i \in \mathbb{R}^N$ is a vector with the ith entry equal to 1 and the other entries equal to 0, and scalars w_{pp} and w_{vp} are defined in (5.12b) and (5.12c), respectively.

Proof. We define the Hamiltonian for UAV i as

$$H_i = \sum_{j=1}^N \frac{\alpha_j r_j}{2} \|u_j\|^2 + \sum_{j=1}^N \lambda_{pj}^T \tilde{B}_p u_j + \sum_{j=1}^N \lambda_{vj}^T \tilde{B}_v u_j$$

where vectors λ_{pi} and λ_{vi} are the Lagrangian multipliers. According to the well-known Pontrya-

gin's minimum principle [71], the necessary conditions for optimality are

$$\dot{s}_{pi} = \frac{\partial H_i}{\partial \lambda_{pi}} = \tilde{B}_p u_i, \quad \dot{s}_{vi} = \frac{\partial H_i}{\partial \lambda_{vi}} = \tilde{B}_v u_i \tag{5.15a}$$

$$\dot{\lambda}_{pi} = -\frac{\partial H_i}{\partial s_{pi}} = 0, \quad \dot{\lambda}_{vi} = -\frac{\partial H_i}{\partial s_{vi}} = 0, \tag{5.15b}$$

$$\lambda_{pi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{pi}(t_f) - s_{pj}(t_f) - \mu_{ij}], \tag{5.15c}$$

$$\lambda_{vi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{vi}(t_f) - s_{vj}(t_f)], \tag{5.15d}$$

$$\frac{\partial H_i}{\partial u_i} = \alpha_i r_i u_i + \tilde{B}_p^T \lambda_{pi} + \tilde{B}_v^T \lambda_{vi} = 0, \quad \frac{\partial^2 H_i}{\partial u_i^2} = \alpha_i r_i > 0. \tag{5.15e}$$

Conditions (5.15b)-(5.15d) indicate that λ_{pi} and λ_{vi} are constant vectors. Substituting them into (5.15e) yields

$$\begin{aligned}
u_i &= -\frac{1}{\alpha_i r_i} \tilde{B}_p^T \lambda_{pi} - \frac{1}{\alpha_i r_i} \tilde{B}_v^T \lambda_{vi} \\
&= -\frac{1}{\alpha_i r_i} \tilde{B}_p^T \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{pi}(t_f) - s_{pj}(t_f) - \mu_{ij}] \\
&\quad - \frac{1}{\alpha_i r_i} \tilde{B}_v^T \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{vi}(t_f) - s_{vj}(t_f)].
\end{aligned} \tag{5.16}$$

Substituting (5.16) into (5.15a) and integrating both sides from 0 to t_f yield

$$\begin{aligned}
& s_{pi}(t_f) + \frac{w_{pp}}{\alpha_i r_i} \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{pi}(t_f) - s_{pj}(t_f)] \\
& + \frac{w_{pv}}{\alpha_i r_i} \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{vi}(t_f) - s_{vj}(t_f)] = s_{pi}(0) + \frac{w_{pp}}{\alpha_i r_i} \mu_i
\end{aligned} \tag{5.17}$$

and

$$s_{vi}(t_f) + \frac{w_{vp}}{\alpha_i r_i} \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{pi}(t_f) - s_{pj}(t_f)]$$

$$+ \frac{w_{vv}}{\alpha_i r_i} \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j)[s_{vi}(t_f) - s_{vj}(t_f)] = s_{vi}(0) + \frac{w_{vp}}{\alpha_i r_i} \mu_i \qquad (5.18)$$

where scalars $w_{pp}, w_{pv}, w_{vp}, w_{vv}$ are defined in (5.12b)-(5.12c) and μ_i is defined in (5.14e). Combining (5.17) and (5.18) and stacking them from $i = 1$ to $i = N$ yield

$$\begin{bmatrix} s_p(t_f) \\ s_v(t_f) \end{bmatrix} = M^{-1} \left\{ \begin{bmatrix} s_p(0) \\ s_v(0) \end{bmatrix} + \left(\begin{bmatrix} w_{pp} \\ w_{vp} \end{bmatrix} \otimes R^{-1} \otimes I_3 \right) \mu \right\} \qquad (5.19)$$

where vectors s_p and s_v are defined in (5.14a), matrix M is defined in (5.11) and invertible according to Lemma 5.1, and vector μ is defined in (5.14d). Therefore, rewriting (5.16) as

$$u_i = -\frac{1}{\alpha_i r_i} F_i \begin{bmatrix} s_p(t_f) \\ s_v(t_f) \end{bmatrix} + \frac{1}{\alpha_i r_i} \tilde{B}_p^T \mu_i \qquad (5.20)$$

where F_i is defined in (5.14b) and substituting (5.19) into (5.20) yields (5.13). Since $s_p(0), s_v(0)$ are in fact functions of the initial state $z(0)$ through (5.6), strategies $u_1^*, \cdots, u_N^*$ in (5.13) form an open-loop Nash equilibrium. $\qquad \square$

Due to the information graph constraint, the following terminal position and velocity estimation law similar to Theorem 3.4 is presented.

Theorem 5.2. *If UAV i updates its vector h_i in continuous time from any initial guess $h_{pi}(0)$ and*

$h_{vi}(0)$ *according to*

$$\begin{bmatrix} \dot{h}_{pi} \\ \dot{h}_{vi} \end{bmatrix} = g \Bigg\{ \begin{bmatrix} s_{pi}(0) \\ s_{vi}(0) \end{bmatrix} + \frac{1}{\alpha_i r_i} \begin{bmatrix} w_{pp} \\ w_{vp} \end{bmatrix} \mu_i - \begin{bmatrix} h_{pi} \\ h_{vi} \end{bmatrix}$$
$$- \frac{1}{\alpha_i r_i} (W \otimes I_3) \sum_{e_{ij} \in \mathcal{E}} (\alpha_i + \alpha_j) \left(\begin{bmatrix} h_{pi} \\ h_{vi} \end{bmatrix} - \begin{bmatrix} h_{pj} \\ h_{vj} \end{bmatrix} \right) \Bigg\} \quad (5.21)$$

where g is a positive scalar, matrix W is defined in (5.12a), and vector μ_i is defined in (5.14e),

then

$$\lim_{\tau \to \infty} \begin{bmatrix} h_p(\tau) \\ h_v(\tau) \end{bmatrix} = \begin{bmatrix} s_p(t_f) \\ s_v(t_f) \end{bmatrix}, \quad (5.22)$$

where $h_p = [h_{p1}^T \ \cdots h_{pN}^T]^T$, $h_v = [h_{v1}^T \ \cdots h_{vN}^T]^T$, and vectors $s_p(t_f)$ and $s_v(t_f)$ are as defined in

(5.19).

Proof. Stacking equation (5.21) from $i = 1$ to $i = N$ yields

$$\begin{bmatrix} \dot{h}_p \\ \dot{h}_v \end{bmatrix} = g \left\{ \begin{bmatrix} s_p(0) \\ s_v(0) \end{bmatrix} - M \begin{bmatrix} h_p \\ h_v \end{bmatrix} + W_\mu \mu \right\}. \quad (5.23)$$

where matrix M is defined in (5.11), matrix W_μ is defined in (5.14c), and vector μ is defined in (5.14d). Since all the eigenvalues of matrix M has positive real parts as shown in Lemma 5.1, matrix $(-M)$ is Hurwitz. Therefore, linear system with respect to $[h_p; h_v]$ in (5.23) is asymptotically stable starting from any initial condition $h(0)$ and will converge to the equilibrium, i.e.,

$$\lim_{\tau \to \infty} \begin{bmatrix} h_p(\tau) \\ h_v(\tau) \end{bmatrix} = M^{-1} \left\{ \begin{bmatrix} s_p(0) \\ s_v(0) \end{bmatrix} + W_\mu \mu \right\}.$$

where the right hand side of the above equation is equal to the vector $[s_p(t_f); s_v(t_f)]$ defined in

(5.19). Therefore, equation (5.22) holds. $\square$

Given the state estimation law in (5.21), an online open-loop Nash strategy design algorithm similar to (3.22) is proposed as follows:

$$u_i^* = -\frac{1}{\alpha_i r_i}\tilde{B}_p^T \sum_{e_{ij}\in\mathcal{E}}(\alpha_i + \alpha_j)(h_{pi} - h_{pj} - \mu_{ij}) - \frac{1}{\alpha_i r_i}\tilde{B}_v^T \sum_{e_{ij}\in\mathcal{E}}(\alpha_i + \alpha_j)(h_{vi} - h_{vj}), \quad (5.24)$$

where h_{pi} and h_{vi} satisfy the equation (5.21) for all $i = 1, \cdots, N$.

For illustration, we apply this online open-loop Nash strategy design algorithm to a five-UAV system. The parameters in [49] are utilized for this simulation: The weight of UAV i is $m_i = 20$kg for all $i = 1, \cdots, N$. The gravity constant is $g = 9.81$kg/m^2. The drag D_i is calculated as follows [78]:

$$D_i = \frac{0.5\rho(V_i - V_{wi})^2 S C_{D0} + 2k_d k_n^2 L^2/g^2}{\rho(V_i - V_{wi})^2 S}$$

where ρ is the atmospheric density and equal to 1.225kg/m^3, V_{wi} is the gust, S is the wing area and equal to 1.37m^2, C_{D0} is the zero-lift drag coefficient and equal to 0.02, k_d is the induced drag coefficient and equal to 0.1, and k_n is the load-factor effectiveness and equal to 1. The gust V_{wi} is modeled as follows [79]:

$$V_{wi} = \bar{V}_{wi} + \delta V_{wi}$$

$$\bar{V}_{wi} = 0.215V_m\log_{10}(h_i) + 0.285V_m$$

where $\bar{V}_{wi}$ is the normal wind shear, V_m is the mean wind speed and equal to 4m/s at the altitude of 80m, and δV_{wi} is the wind gust turbulence on UAV i and assumed to be a Gaussian random variable with zero mean and a standard deviation equal to $0.09V_m$. The real control inputs of the UAV have

the following constraints: $T_i < 125N$, $-294.3N < L_i < 392.4N$, and $-80° \leq \phi_i \leq 80°$ for all $i = 1, \cdots, N$. We assume that the UAVs are trying to form a desired V-shape in the same altitude shown in Figure 5.2 and the underlying undirected information graph is also shown in this figure.
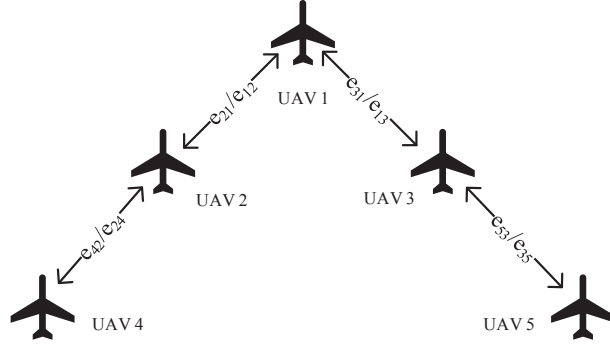

Figure 5.2: V-shape formation and information graph

Hence, UAV 1 will act as a reference for the other UAVs. The corresponding graph Laplacian matrix is

$$\mathcal{L} = \begin{bmatrix} 2\alpha_1 + \alpha_2 + \alpha_3 & -\alpha_1 - \alpha_2 & -\alpha_1 - \alpha_3 & 0 & 0 \\ -\alpha_1 - \alpha_2 & \alpha_1 + \alpha_4 + 2\alpha_2 & 0 & -\alpha_4 - \alpha_2 & 0 \\ -\alpha_1 - \alpha_3 & 0 & \alpha_1 + \alpha_5 + 2\alpha_3 & 0 & -\alpha_3 - \alpha_5 \\ 0 & -\alpha_2 - \alpha_4 & 0 & \alpha_2 + \alpha_4 & 0 \\ 0 & 0 & -\alpha_3 - \alpha_5 & 0 & \alpha_3 + \alpha_5 \end{bmatrix}.$$

where $0 \leq \alpha_1, \cdots, \alpha_5 \leq 1$ are the convex parameters and $\sum_{i=j}^{5} \alpha_j = 1$. The desired offset vectors of the formation among the UAVs are

$$\mu_{21} = \begin{bmatrix} -100 \\ -100 \\ 0 \end{bmatrix} \text{m}, \mu_{31} = \begin{bmatrix} 100 \\ -100 \\ 0 \end{bmatrix} \text{m}, \mu_{42} = \begin{bmatrix} -100 \\ -100 \\ 0 \end{bmatrix} \text{m}, \mu_{53} = \begin{bmatrix} 100 \\ -100 \\ 0 \end{bmatrix} \text{m}.$$

The initial positions of the UAVs are

$$p_1(0) = \begin{bmatrix} 0 \\ 0 \\ 90 \end{bmatrix} \text{m}, \ p_2(0) = \begin{bmatrix} -80 \\ 0 \\ 80 \end{bmatrix} \text{m}, \ p_3(0) = \begin{bmatrix} 90 \\ 0 \\ 70 \end{bmatrix} \text{m}, \ p_4(0) = \begin{bmatrix} -120 \\ 0 \\ 60 \end{bmatrix} \text{m}, \ p_5(0) = \begin{bmatrix} 150 \\ 0 \\ 65 \end{bmatrix} \text{m}.$$

The initial velocities of the UAVs are

$$v_1(0) = \begin{bmatrix} 0 \\ 50 \\ 0 \end{bmatrix} \text{m/s}, \ v_2(0) = \begin{bmatrix} 0 \\ 60 \\ 0 \end{bmatrix} \text{m/s}, \ v_3(0) = \begin{bmatrix} 0 \\ 40 \\ 0 \end{bmatrix} \text{m/s}, \ v_4(0) = \begin{bmatrix} 0 \\ 65 \\ 0 \end{bmatrix} \text{m/s}, \ v_5(0) = \begin{bmatrix} 0 \\ 45 \\ 0 \end{bmatrix} \text{m/s}.$$

The five UAVs' performance indices are given by (5.5) with $t_f = 30$ and $r_i = 1$ for all $i = 1, \cdots, 5$. With $\alpha_i = 0.2$ for all $i = 1, \cdots, 5$, the the the UAVs' trajectories derived using the online open-loop Nash strategy design algorithm (5.24) are shown in Figure 5.3. The left plot shows the trajectories in 3-dimensional space and the right plot shows the trajectories in $x - y$ plane. In the figure, the circles indicate the UAVs' initial positions and the triangles indicate the UAVs' terminal positions.
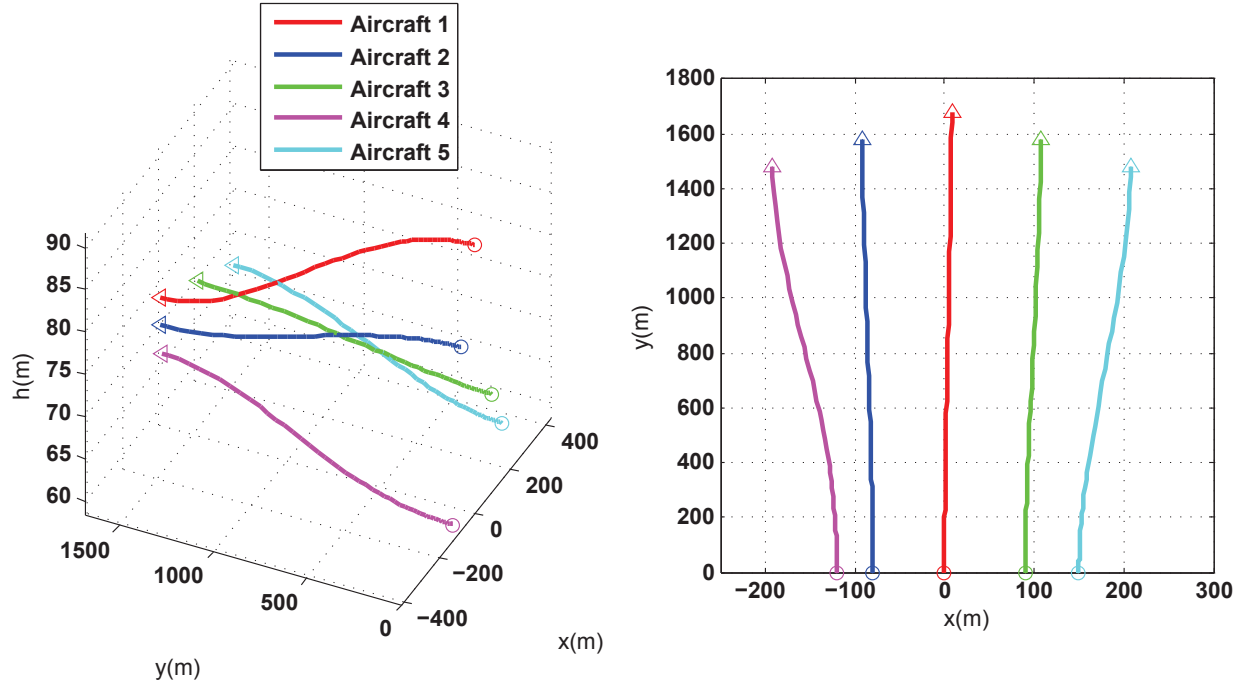
Figure 5.3: UAVs' trajectories in 3D and $x - y$ plane

Clearly, the UAVs' positions form the desired V-formation at the terminal time. For illustrative purpose, the UAVs' trajectories on x axis, y axis, and h axis are shown independently in Figure 5.4. Moreover, the UAVs' velocities in the three axis and three real control inputs obtained according to the relationship (5.3) are also shown in Figure 5.4. All of real control inputs are within the specified constraints.
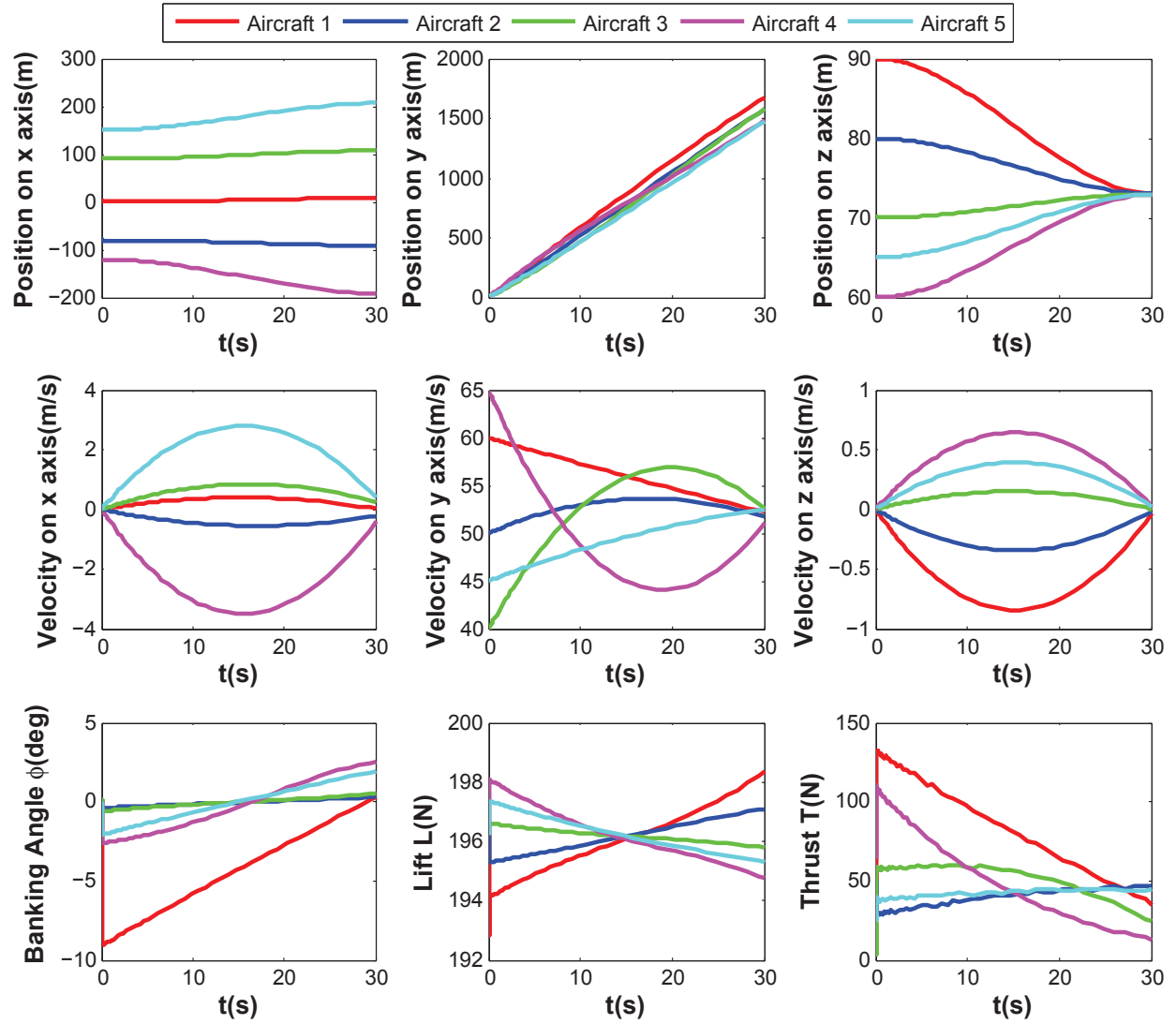
Figure 5.4: UAVs' positions, velocities, and real control inputs

5.2 Multi-Pursuer Single-Evader Differential Game with Limited Observations

In this section, we consider a differential game over a finite time horizon in which only the evader is assumed to have global sensing capability which allows it to observe all the pursuers at all times. Each pursuer, on the other hand, has a limited sensing capability which allows it to observe the evader and/or other pursuers only if they fall within its sensing range. A practical example of such a situation occurs when a well-equipped UAV with a very wide range of sensing capability must evade several (possibly a large number of) weakly-equipped pursuing UAVs. In what follows, we derive the feedback Nash strategies for both the pursuers and evader using the best achievable performance indices based approach.

We define the following displacement vector z_i between pursuer i and the evader e as shown in Figure 5.5

$$z_i = x_e - x_i \qquad \forall i = 1, \cdots, N. \tag{5.25}$$

where $x_e \in \mathbb{R}^n$ is the evader's position vector and $x_i \in \mathbb{R}^n$ is pursuer i's position vector.
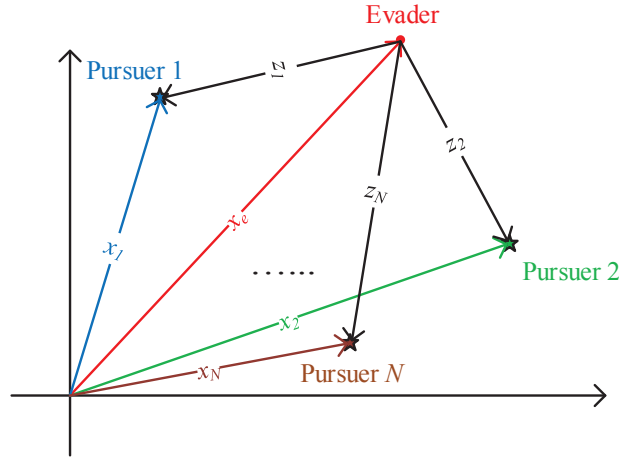

Figure 5.5: Displacement vectors

We assume that a collective objective of the pursuers is to minimize the sum of the weighted distances between the evader and themselves at a terminal time $t_f > 0$ while at the same time minimizing these distances and their control efforts over the time interval $[0, t_f]$. Hence, the group of pursuers tries to minimize the following performance index:

$$J_p = \sum_{j=1}^{N} \frac{f_{p_j}}{2} \|z_j(t_f)\|^2 + \int_0^{t_f} \sum_{j=1}^{N} \left(\frac{q_{p_j}}{2} \|z_j\|^2 + \frac{r_j}{2} \|u_j\|^2 \right) dt, \qquad (5.26)$$

where u_j is pursuer j's velocity control input, $\|\cdot\|$ is the Euclidean norm, and scalars f_{p_j}, q_{p_j}, and r_j are positive weights for $j = 1, \cdots, N$. On the other hand, we assume that the evader's objective is to maximize the sum of the weighted terminal distances between the pursuers and itself while at the same time maximizing these distances and minimizing its control effort over the time interval $[0, t_f]$. Hence, the evader will try to minimize the performance index:

$$J_e = - \sum_{j=1}^{N} \frac{f_{e_j}}{2} \|z_j(t_f)\|^2 + \int_0^{t_f} \sum_{j=1}^{N} \left(-\frac{q_{e_j}}{2} \|z_j\|^2 \right) + \frac{r_e}{2} \|u_e\|^2 dt, \qquad (5.27)$$

where u_e is the evader's velocity control input and scalars f_{e_j}, q_{e_j}, and r_e are positive weights for $j = 1, \cdots, N$. To express the system dynamics more compactly, we define the vector $z = [z_1^T; \cdots z_N^T]^T$ which, along with (5.25), yields

$$\dot{z} = B_e u_e + B_p u_p. \qquad (5.28)$$

where matrix $B_e = \mathbf{1}_N \otimes I_n$, $\mathbf{1}_N \in \mathbb{R}^{N \times 1}$ is a vector with all the entries equal to 1, $u_p = [u_1^T \ \cdots \ u_N^T]^T$, $B_p = -I_N \otimes I_n$. The performance indices (5.26) and (5.27) can be rewritten

as

$$J_p = \frac{1}{2}\|z(t_f)\|_{F_p}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_p}^2 + \|u_p\|_{R_p}^2)\mathrm{d}t, \tag{5.29a}$$

$$J_e = \frac{1}{2}\|z(t_f)\|_{F_e}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_e}^2 + \|u_e\|_{R_e}^2)\mathrm{d}t, \tag{5.29b}$$

where $\|z\|_F^2 = z^T F z$ and

$$F_p = \mathrm{diag}\{f_{p_1}, \cdots, f_{p_N}\} \otimes I_n, \qquad F_e = -\mathrm{diag}\{f_{e_1}, \cdots, f_{e_N}\} \otimes I_n$$

$$Q_p = \mathrm{diag}\{q_{p_1}, \cdots, q_{p_N}\} \otimes I_n, \qquad Q_e = -\mathrm{diag}\{q_{e_1}, \cdots, q_{e_N}\} \otimes I_n$$

$$R_p = \mathrm{diag}\{r_1, \cdots, r_N\} \otimes I_n,$$

and "diag" stands for "diagonal matrix". Hence, given the system dynamics in (5.28) and performance indices in (5.29), a differential nonzero-sum game between the group of pursuers and the evader is formed. To accurately model the sensing capabilities and limited observations of the pursuers, we assume that pursuer i has a sensing range defined by a sensing radius $r_i > 0$. If the Euclidean distance between pursuer i and the evader is less than or equal to r_i, that is, $\|x_i - x_e\| \le r_i$, then pursuer i is able to observe the evader, otherwise, pursuer i cannot observe the evader. Consequently, we define a binary scalar $h_i(t)$ to represent pursuer i's ability to observe the evader at time t as follows:

$$h_i(t) = \begin{cases} 1 & \text{if } \|x_i(t) - x_e(t)\| \le r_i \\ 0 & \text{if } \|x_i(t) - x_e(t)\| > r_i \end{cases}. \tag{5.30}$$

Similarly, if the the Euclidean distance between pursuer i and pursuer j is less than or equal to r_i, that is, $\|x_i - x_j\| \le r_i$, then pursuer i is able to observe pursuer j, otherwise, pursuer i cannot observe pursuer j. Consequently, we can use an unweighted Laplacian matrix, $L(t) = [L_{ij}(t)] \in$

$\mathbb{R}^{N \times N}$, similar to (2.4) to described the observations among the pursuers at every instant of time t, where

$$L_{ij}(t) = \begin{cases} -1 & \text{if } \|x_i(t) - x_j(t)\| \leq r_i \text{ for } j \neq i \\ 0 & \text{if } \|x_i(t) - x_j(t)\| > r_i \text{ for } j \neq i \\ -\displaystyle\sum_{l=1,l \neq i}^{N} L_{il} & \text{if } j = i \end{cases} \tag{5.31}$$

for $i, j = 1, \cdots, N$.

For the formulated pursuit-evasion game with limited observations, the following practical issue needs to be noticed: Although the evader has sufficiently wide observation range to observe all the pursuers' positions at every instant of time, it really has no information on the individual pursuers' observation radii $r_1, \cdots, r_N$ or how these pursuers obverse each other among themselves. Therefore, we assume that during the game process, the evader has no knowledge of the existence of limited observations among the pursuers and the overall information topology. On the other hand, for the pursuers, we assume that all of them are aware of their limited observation capabilities as well as the evader's global observation capability.

Given the formulated pursuit-evasion game problem, every player is able to solve for the Nash equilibrium using the well-known Riccati equation approach, however, only the evader who observes all the pursuers can implement this Nash strategy. According to [7], for the game defined by system (5.28) and performance indices (5.29), the classical linear feedback Nash strategies are

$$u_p^* = -R_p^{-1} B_p^T P_p z \tag{5.32a}$$

$$u_e^* = -R_e^{-1} B_e^T P_e z, \tag{5.32b}$$

where matrices P_p and P_e are solutions to the coupled differential Riccati equations

$$\dot{P}_p + Q_p - P_p B_p R_p^{-1} B_p^T P_p - P_p B_e R_e^{-1} B_e^T P_e - P_e B_e R_e^{-1} B_e^T P_p = 0 \tag{5.33a}$$

$$\dot{P}_e + Q_e - P_e B_p R_p^{-1} B_p^T P_p - P_p B_p R_p^{-1} B_p^T P_e - P_e B_e R_e^{-1} B_e^T P_e = 0. \tag{5.33b}$$

with boundary condition $P_p(t_f) = F_p$ and $P_e(t_f) = F_e$. The expressions for the evader's Nash strategy in (5.32b) is indeed linear feedback controls of the global state vector z. Since the evader has no knowledge of the existence of the pursuers' limited observations, it naturally implements the feedback Nash strategy (5.32b) as its control input during the game, assuming that the pursuers are implementing strategy (5.32a). From the pursuers' perspective, since each of them has limited observation, there is no way for them to implement their Nash strategy described in (5.32a). Therefore, a Nash strategy design approach must be proposed for the group of pursuers to accommodate their limited observations constraint while at the same time maintaining a Nash equilibrium with the evader's strategy (5.32b). First of all, the pursuers' admissible control u_p needs to be properly structured in order to fit into the limited observation constraint that each pursuer must operate under. Toward this end, we propose the following structured feedback strategies for the pursuers:

$$u_i^s = \underbrace{h_i(t) K_{ie}(t) z_i(t)}_{a} + \underbrace{K_{ip}(t) \sum_{j=1}^{N} L_{ij}(t) z_j(t)}_{b} \quad \forall i = 1, \cdots, N \tag{5.34}$$

where scalar $h_i(t)$ is defined in (5.30), scalar $L_{ij}(t)$ is defined in (5.31), matrices $K_{ie} \in \mathbb{R}^{n \times n}$ and $K_{ip} \in \mathbb{R}^{n \times n}$ are feedback gains to be determined. Term a in (5.34) represents a control component of pursuer i to chase the evader directly if it observes the evader. Term b in (5.34) is known as a cooperative control component, that is, a feedback control of the difference between the position of pursuer i and those of the pursuers that it observes. The expression of pursuer i's control in (5.34) means that when pursuer i is able to observe the evader (i.e. when $h_i(t) = 1$), it will chase the

evader while at the same time it follows the nearby pursuers that it can observe. When pursuer i is unable to observe the evader (i.e. when $h_i(t) = 0$), it has no choice but to merely follow the nearby pursuers. The control expression in (5.34) can be rewritten using the more compact notation as

$$u_i^s = K_{ie}[0 \ \cdots \ 0 \ (h_i I_n) \ 0 \ \cdots \ 0]z + K_{ip}[(L_{i1} I_n) \ \cdots \ (L_{iN} I_n)]z \triangleq M_i C_i z, \qquad (5.35)$$

where $M_i = [K_{ie} \ K_{ip}] \in \mathbb{R}^{n \times 2n}$ and

$$C_i = \begin{bmatrix} 0 & \cdots & 0 & (h_i I_n) & 0 & \cdots & 0 \\ (L_{i1} I_n) & \cdots & & & \cdots & (L_{iN} I_n) \end{bmatrix},$$

where h_i and L_{ij} are defined in (5.30) and (5.31). Therefore, the pursuers' control vector $u_p^s = [(u_1^s)^T \ \cdots \ (u_N^s)^T]^T$ can be written as $u_p^s = M_p z$, where

$$M_p = [(M_1 C_1)^T \ \cdots \ (M_N C_N)^T]^T. \qquad (5.36)$$

The problem now reduces to finding a set of matrices $M_1^*, \cdots, M_N^*$ such that feedback gain $M_p^* = [(M_1^* C_1)^T \ \cdots \ (M_N^* C_N)^T]^T$ and the resulting pursuers' strategy

$$u_p^{s*} = M_p^* z \qquad (5.37)$$

can still form a Nash equilibrium with the evader's strategy u_e^* in (5.32b). As we mentioned in Chapter 4, using the optimal output feedback based approach, all the players' structured controls and the corresponding feedback gains need to be simultaneously parameterized and optimized with respect to the given set of performance indices to obtain the Nash equilibrium. This cannot be implemented in our game setup since as mentioned earlier, the evader will be implementing the Nash strategy (5.32b) and hence it is not possible to simultaneously parameterize and optimize it

along with the strategy of the pursuers to form a Nash equilibrium. Therefore, the best achievable performance indices based approach is utilized to design Nash strategies for the pursuers. The following result similar to Theorem 4.5 is presented.

Theorem 5.3. *For the pursuit-evasion game described by system dynamics (5.28) and performance indices (5.29), for an arbitrary set of matrices $M_1, \cdots, M_N$, the strategies u_e^* in (5.32b) and $u_p^s = M_p z$ form a Nash equilibrium*

$$J_p^s(u_p^s, u_e^*) \leq J_p^s(u_p, u_e^*), \quad \forall u_p \in U_p \tag{5.38a}$$

$$J_e^s(u_p^s, u_e^*) \leq J_e^s(u_p^s, u_e), \quad \forall u_e \in U_e, \tag{5.38b}$$

with respect to performance indices

$$J_p^s = \frac{1}{2}\|z(t_f)\|_{F_p}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_p^s}^2 - u_p^T \Gamma z - z^T \Gamma^T u_p + \|u_p\|_{R_p}^2)dt, \tag{5.39a}$$

$$J_e^s = \frac{1}{2}\|z(t_f)\|_{F_e}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_e^s}^2 + \|u_e\|_{R_e}^2)dt, \tag{5.39b}$$

where

$$Q_p^s = M_p^T R_p M_p - P_p B_p R_p^{-1} B_p^T P_p + Q_p \tag{5.40a}$$

$$\Gamma = B_p^T P_p + R_p M_p, \tag{5.40b}$$

$$Q_e^s = -P_e B_p R_p^{-1}(R_p M_p + B_p^T P_p) - (R_p M_p + B_p^T P_p)^T R_p^{-1} B_p^T P_e + Q_e. \tag{5.40c}$$

matrices P_p and P_e are the solutions to (5.33) and matrix M_p is as given in (5.36).

Proof. Consider Lyapunov functions

$$V_p = \frac{1}{2}z^T P_p z \quad \text{and} \quad V_e = \frac{1}{2}z^T P_e z. \tag{5.41}$$

Differentiating V_p in (5.41) with respect to t and integrating it from 0 to t_f yield

$$V_p(t_f) - V_p(0) = \frac{1}{2} \int_0^{t_f} \Big[-\|z\|_{Q_p^s}^2 - \|u\|_{R_p}^2 + u_p^T \Gamma z + z^T \Gamma^T u_p$$
$$+ 2z^T P_p B_e (u_e + R_e^{-1} B_e^T P_e z) + \|u_p - M_p z\|_{R_p}^2 \Big] \mathrm{d}t.$$

Hence,

$$J_p^s = V_p(0) + \int_0^{t_f} \frac{1}{2} \|u_p - M_p z\|_{R_p}^2 + z^T P_s B_e (u_e + R_e^{-1} B_e^T P_e z) \mathrm{d}t, \qquad (5.42)$$

where J_p^s is defined in (5.39a). Similarly, we can show that

$$J_e^s = V_e(0) + \int_0^{t_f} \frac{1}{2} \|u_e + R_e^{-1} B_e^T P_e z\|_{R_e}^2 - z^T P_e B_p (u_p - M_p z) \mathrm{d}t, \qquad (5.43)$$

where J_e^s is defined in (5.39b). Since R_p and R_e are positive definite, it is obvious from (5.42) and (5.43) that given performance indices defined in (5.39), the inequalities in (2.8) holds for $u_p = M_p z$ and $u_e^* = -R_e^{-1} B_e^T P_e$. Hence, strategies (5.32b) and (5.37) form a Nash equilibrium with respect to performance indices in (5.39). Clearly, if M_p can be written as $M_p = -R_p^{-1} B_p^T P_p$, then J_p^s in (5.39a) becomes identical to (5.29a) and J_e^s in (5.39b) becomes identical to (5.29b). $\qquad \square$

Therefore, according to the definition of best achievable performance index in (4.1), to find the optimal matrix $M_p^*(t)$ corresponding to the best achievable performance indices, we need to solve a multi-objective optimization problem of minimizing $\|Q_p^s - Q_p\|_f^2$, $\|S\|_f^2$, and $\|Q_e^s - Q_e\|_f^2$ simultaneously. Hence, we let $H(t)$ in (4.19) be

$$H(t) = \beta_1 \|Q_p^s - Q_p\|_f^2 + \beta_2 \|S\|_f^2 + \beta_3 \|Q_e^s - Q_e\|_f^2$$
$$= \beta_1 \mathrm{Tr}[(Q_p^s - Q_p)^2] + \beta_2 \mathrm{Tr}(S^T S) + \beta_3 \mathrm{Tr}[(Q_e^s - Q_e)^2] \qquad (5.44)$$

where $0 < \beta_j < 1$ for $j = 1, 2, 3$, and $\sum_{j=1}^{3} \beta_j = 1$. Since M_p is as defined in (5.36), the minimization in (4.20) is actually done with respect to $M_1(t), \cdots, M_N(t)$. With the gradient of $H(t)$ with respect to $M_1(t), \cdots, M_N(t)$ expressed as follows:

$$\nabla_{M_i} H = (d_i^T \otimes I_n)[4\alpha_1 R_p M_p (Q_p^s - Q_p) + 2\alpha_2 R_p S - 4\alpha_3 B_p^T P_e (Q_e^s - Q_e)]C_i^T. \qquad (5.45)$$

for all $i = 1, \cdots, N$, where $d_i \in \mathbb{R}^N$ is a vector with the ith entry equal to 1 and the other entries equal to 0, a gradient based iterative algorithms similar to Algorithm 4.2 can be adopted to find matrices $M_1^*(t), \cdots, M_N^*(t)$. Also note that by varying the coefficients $\beta_1, \beta_2, \beta_3$ in (5.44), a noninferior set of the solutions can be generated. An appropriate choice of these coefficients can be made to place a desired emphasis on the importance of minimizing each of the three terms in (5.44) as compared to the other two.

For illustrative purpose, let us consider a three-pursuer single-evader differential game taking place in a planar environment and defined over a time interval $[0, 3]$. Suppose that $x_i = [x_{i1}^T\ x_{i2}^T]^T \in \mathbb{R}^2$ represents player i's position and $u_i = [u_{i1}^T\ u_{i2}^T]^T \in \mathbb{R}^2$ represents player i's velocity control. Hence, in equation (2.2), we have

$$B_e = \begin{bmatrix} I_2 \\ I_2 \\ I_2 \end{bmatrix} \quad \text{and} \quad B_p = -\begin{bmatrix} I_2 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \end{bmatrix}.$$

The performance indices are given by (5.29) with $t_f = 3$, $F_p = Q_p = qI_6$, $R_p = I_6$, $F_e = Q_e = I_6$, and $R_e = I_2$, where q is a positive scalar that can be varied to analyze different scenarios. As shown in Figure 5.6, we assume that the pursuers' initial positions are $x_1(0) = (-3, 0)$, $x_2(0) = (3, 0)$, $x_3(0) = (5, 1)$, the evader's initial position is $x_e(0) = (0, 1)$, and the pursuers' sensing radii are the same and equal to 4. Clearly, at $t = 0$, pursuer 1 can only observe the evader, pursuer 2 can
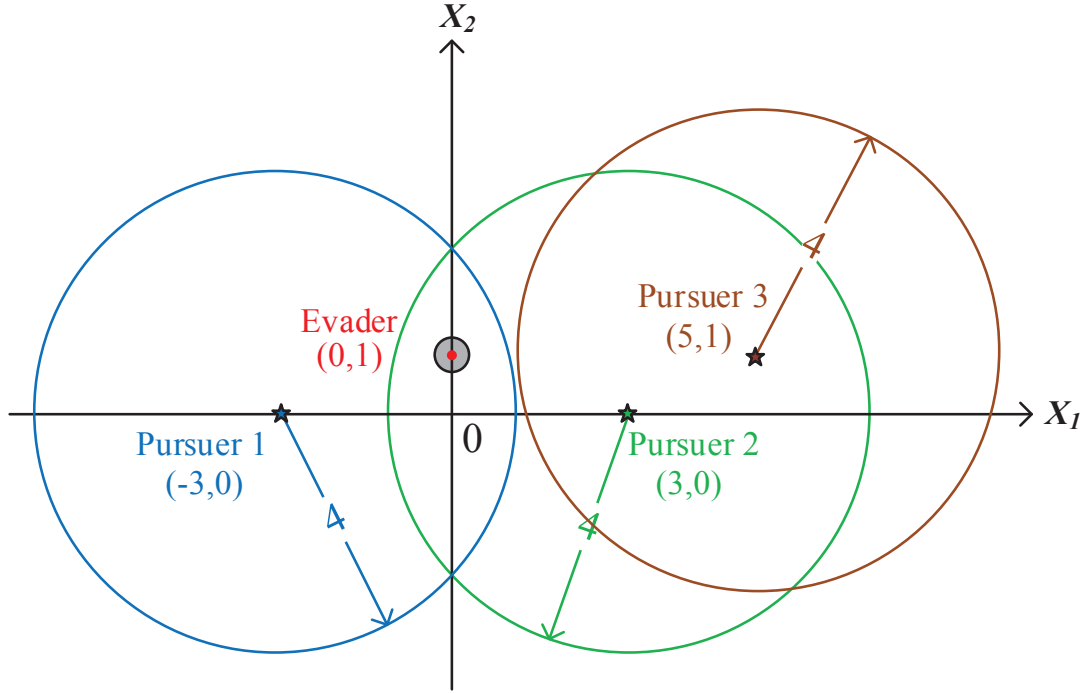
Figure 5.6: Initial positions of three pursuers and single evader

observe the evader and pursers 3, and pursuer 3 can only observe pursuer 2. Further, we assume that the evader is captured if the minimum distance between the pursuers and evader is less than a capture radius $\sigma = 0.1$, which is shown as a light black circle centered at the evader in Figure 5.6. In this example, we will consider two different scenarios:

- **Scenarios 1**: $q = 1$. Pursuers put equal emphasis on minimizing their distances to the evader and minimizing their control effort.

- **Scenario 2**: $q = 5$. Pursuers put more emphasis on minimizing their distances to the evader than on minimizing their control effort.

Evader's Strategy: The Evader solves the coupled differential Riccati equations (5.33) and im-

plements the corresponding Nash strategy. This yields solutions P_p and P_e in the following form:

$$
P_p = \begin{bmatrix} P_{p1} & P_{p2} & P_{p2} \\ P_{p2} & P_{p1} & P_{p2} \\ P_{p2} & P_{p2} & P_{p1} \end{bmatrix} \otimes I_2, \quad P_e = \begin{bmatrix} P_{e1} & P_{e2} & P_{e2} \\ P_{e2} & P_{e1} & P_{e2} \\ P_{e2} & P_{e2} & P_{e1} \end{bmatrix} \otimes I_2
$$

where the plots of $P_{p1}(t)$, $P_{p2}(t)$, $P_{e1}(t)$, and $P_{e2}(t)$ for both scenarios are shown in Figure 5.7.

Hence, the evader's feedback Nash strategies (5.32b) in terms of z_1, z_2, z_3 can be expressed as

$$
u_e^* = (P_{e1} + 2P_{e2})(z_1 + z_2 + z_3).
$$
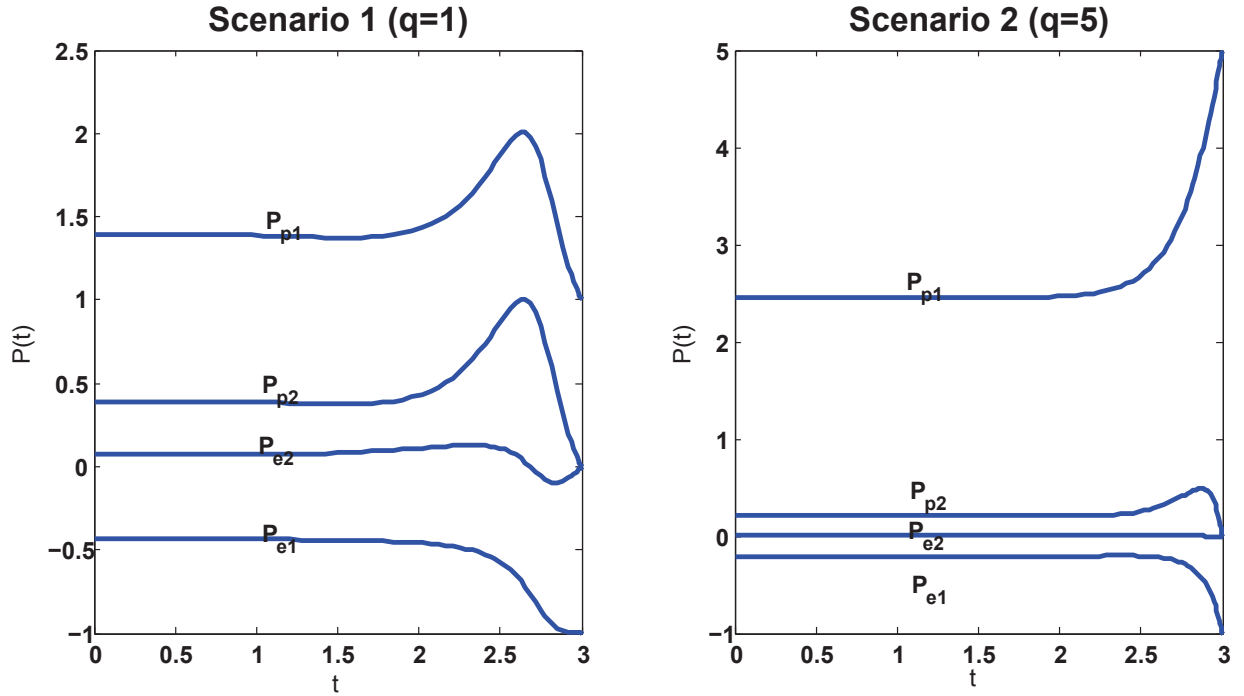
Figure 5.7: Plots of $P_{p1}(t)$, $P_{p2}(t)$, $P_{e1}(t)$, and $P_{e2}(t)$

Pursuers' Strategy: To derive the pursuers' strategy, we assume that for implementation purpose, the pursuers perform sensing only at discrete instants of time $t_0, t_1, \cdots, t_{299}$, where $t_0 = 0$, $t_{300} = t_f = 3$. Since $t_j - t_{j-1} = 0.01$ is quite small for all $j = 1, \cdots, 300$, we assume that the observations among the players can be regarded to be constant within such a small time interval $(t_j - t_{j-1})$. We also assume that the pursuers will carry out the proposed best achievable performance indices approach with the following (arbitrary) choice of coefficients in (5.44): $\alpha_1 = 1/4$, $\alpha_2 = 1/2$, and $\alpha_3 = 1/4$.

Scenario 1: In this scenario, the motion trajectories of the pursuers and evader over time are shown in Figure 5.8. The distances between the pursuers and evader over time are shown in Figure 5.9 where the capture radius $\sigma = 0.1$ is shown in terms of a dashed black horizontal line. Clearly, in this scenario, none of the pursuer is able to capture the evader when the final time $t_f = 3$ is reached. Furthermore, the change in the observations among the players is reflected in the changes of $h_i(t)$ in (5.30) and Laplacian matrix with $L_{ij}(t)$ defined in (5.31).
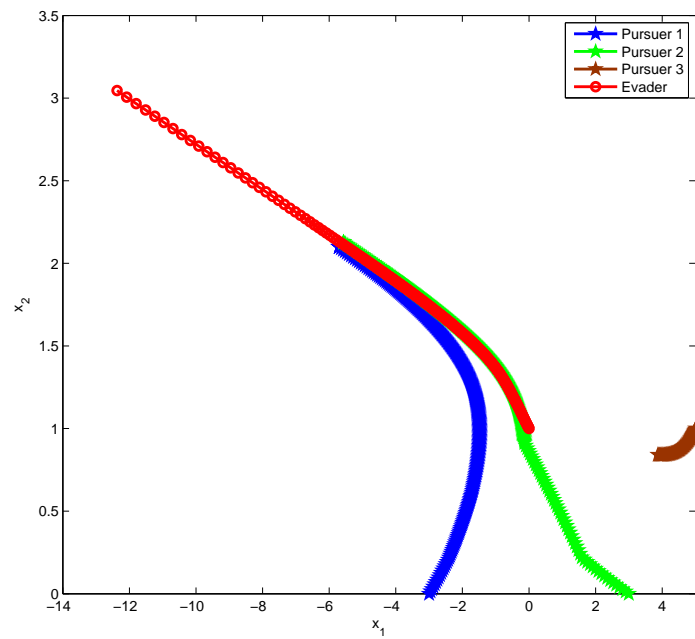
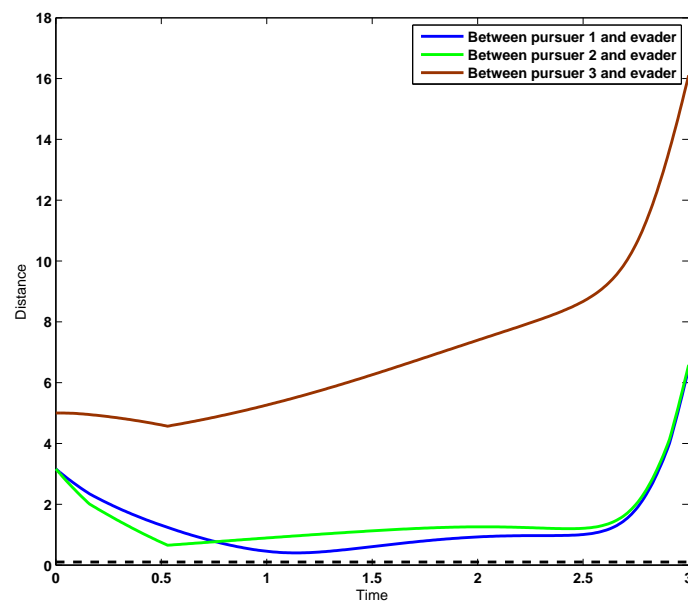Figure 5.8:   Motion trajectories of the pursuers and evader for Scenario 1

Figure 5.9:   Distances between the pursuers and evader for Scenario 1

In this scenario, the values of $h_1(t), h_2(t), h_3(t)$ and the value of the Laplacian matrix $L(t)$ have changed as follows:

$$h_1(t) = \begin{cases} 1 & 0 \leq t \leq 2.91 \\ 0 & 2.91 < t \leq 3 \end{cases}$$

$$h_2(t) = \begin{cases} 1 & 0 \leq t \leq 2.91 \\ 0 & 2.91 < t \leq 3 \end{cases} \quad \text{and} \quad L(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0 \leq t \leq 0.16 \\[2em] \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0.16 < t \leq 0.52 \\[2em] \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & 0.52 < t \leq 3 \end{cases}$$

$$h_3(t) = 0 \quad 0 < t \leq 3$$

The change in $h_i(t)$ means that pursuers 1 and 2 lose observation of the evader after $t = 2.91$ while pursuer 3 was never able to observe the evader for the entire game. The change in the Laplacian matrix essentially means that only pursuers 2 and 3 can observe each other for $t \in [0, 0.16]$, pursuer 2 can observe pursuers 1 and 3 for $t \in (0.16, 0.52]$ while pursuers 1 and 3 cannot observe each other at this time interval, and only pursuers 1 and 2 can observe each other for the rest time $t \in (0.52, 3]$.

Scenario 2: In this scenario, the motion trajectories of the pursuers and evader are shown in Figure 5.10. The distances between the pursuers and evader are shown in Figure 5.11 where the the capture radius $\sigma = 0.1$ is shown in terms of a dashed black horizontal line. Clearly, in this scenario, pursuer 2 is the first one to capture the evader at $t = 1.2$.
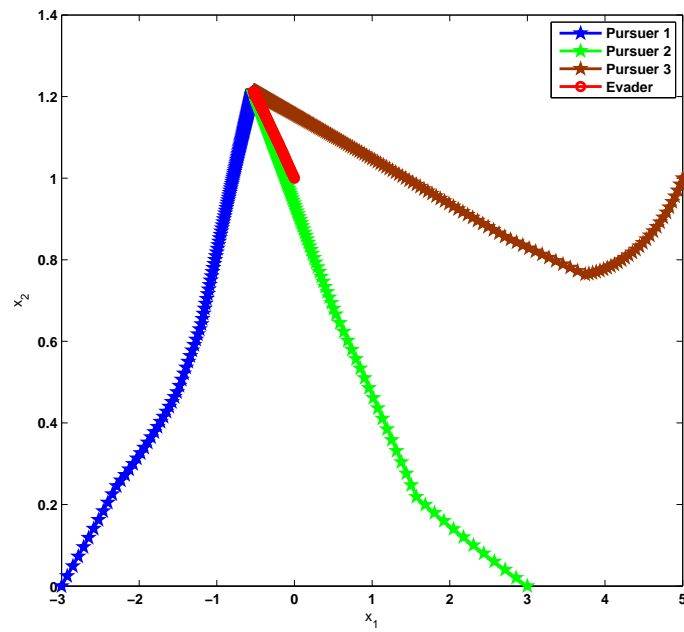
Figure 5.10:   Motion trajectories of the pursuers and evader for Scenario 2
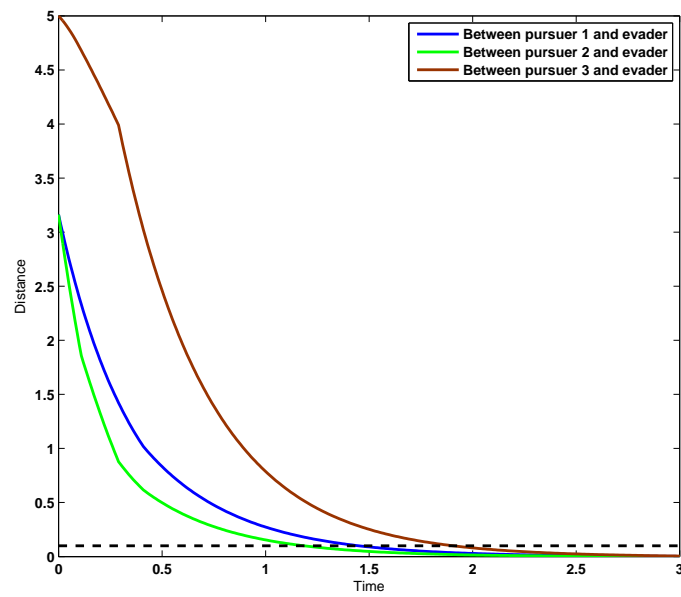

Figure 5.11:   Distances between the pursuers and evader for Scenario 2

During the entire game, the values of $h_1(t), h_2(t), h_3(t)$ and the value of the Laplacian matrix $L(t)$ have changed as follows

$$h_1(t) = 1 \quad 0 < t \leq 3$$

$$h_2(t) = 1 \quad 0 < t \leq 3$$

$$h_3(t) = \begin{cases} 0 & 0 \leq t \leq 0.29 \\ 1 & 0.29 < t \leq 3 \end{cases}$$

$$\text{and} \quad L(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0 \leq t \leq 0.11 \\[2em] \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0.11 < t \leq 0.41 \\[2em] \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} & 0.41 < t \leq 3 \end{cases}$$

The change of $h_i(t)$ means that after $t = 0.29$, all the pursuers are able to observe the evader. The change of the Laplacian matrix means that only pursuers 2 and 3 can observe each other for $t \in [0, 0.11]$, pursuer 2 can observe pursuers 1 and 3 for $t \in (0.11, 0.41]$ while pursuers 1 and 3 cannot observe each other at this time interval, and all the pursuers are able to observe each other for the rest time $t \in (0.41, 3]$.

It would be interesting to determine a critical value q_c of q which separates the escape and capture regions of the evader. That is, if $q < q_c$, the evader escapes and if $q \geq q_c$ the evader is captured at a time instant $t \in [0, 3]$. For this game, the critical value of q has been determined to be $q_c = 1.38$. Figure 5.12 shows the motion trajectories of the pursuers and evader when $q = q_c = 1.38$. Figure 5.13 shows the distances between the pursuers and evader when $q = q_c = 1.38$, where the capture time occurs at $t = 1.55$.
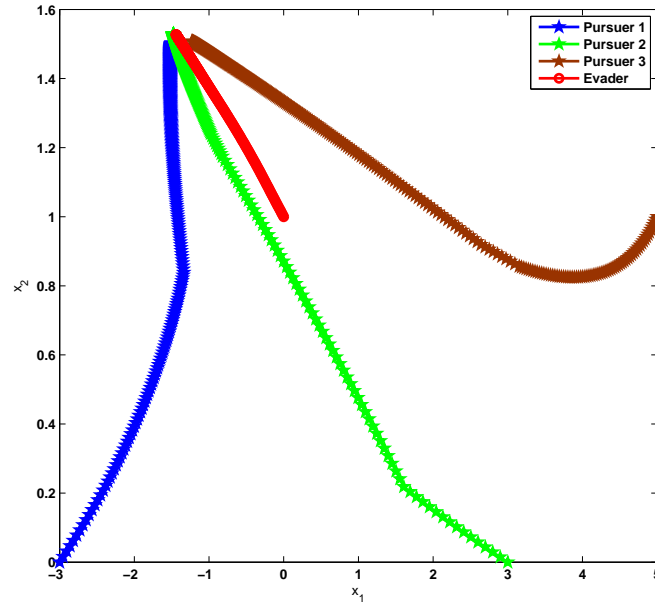
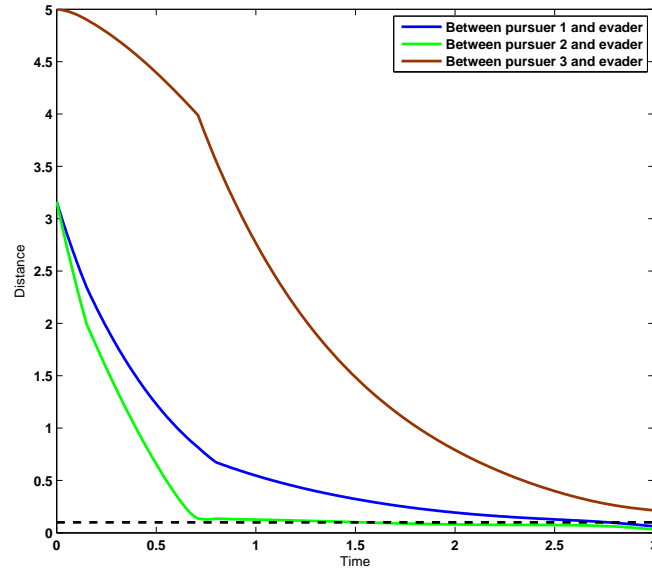Figure 5.12: Trajectories of the pursuers and evader when $q = q_c = 1.38$

Figure 5.13: Distances between the pursuers and evader when $q = q_c = 1.38$

CHAPTER 6: CONCLUSION

This dissertation focuses on the Nash strategy and noninferior strategy designs for linear quadratic differential games in the multi-agent system under distributed open-loop and feedback information structures. We first introduced the basic concepts in the game theory and multi-agent control systems, reviewed the existing results in these fields, raised the motivation of this research, and defined the scope of this dissertation. As the main results, we proposed novel open-loop and feedback game strategy design approaches to overcome the conventional approaches' incapabilities in dealing with the distributed information constraint. The contributions of this dissertation can be addressed as follows:

1. In terms of the open-loop strategy design, the proposed approach integrates a distributed state estimation algorithm into the classical open-loop game strategy.

    – The proposed approach can be carried out in a distributed manner where every agent is able to implement it by exchanging the state estimates with other agents according to the information graph.

    – The proposed approach renders approximate strategies of the original open-loop Nash or noninferior strategies which can only be implemented under global information and these approximate strategies can be made arbitrarily close to the original open-loop Nash or noninferior strategies.

2. In terms of the feedback strategy design, the proposed approach is based on the concept of best achievable performance indices.

    – The proposed approach renders structured strategies which has structured feedback gain matrices that conform to the information graph constraint.

– Compared with the classical output feedback optimal approach, the proposed approach renders strategies that are independent on the initial state of the system and does not requires all the agents to parameterize and optimize the feedback gain matrices simultaneously.

Two illustrative application examples on an unmanned aerial vehicle formation control problem and a multi-pursuer single-evader differential game problem with limited observations were solved and the simulation results corresponding to different scenarios are presented.

With the already obtained results, the future research can be carried out in the following possible directions:

For the open-loop design, the possible directions are as follows. First, applying the idea in the proposed open-loop strategy design approach to feedback strategy design will be significant because if the feedback strategy design approach is obtained, then the problems of more realistic importance including the differential games for multi-agent systems under time-varying information graph can be tackled. Second, the proposed approach can be successfully implemented if the convergence condition of the state estimation law is valid, which is the case for most of the consensus problems. Therefore, exploring the condition in more details or finding conditions where the condition always holds will be interesting and very important.

For the feedback design, the best achievable performance indices approach requires an authority to carry out the computing algorithm with the knowledge of the overall information topology and all the agents' information and then distribute the resulting game strategy to each and every agent. In order to have an better adaptation to the time-varying information graph, an approximate approach to the proposed approach that requires less information need be proposed.

APPENDIX : PAPERS PUBLISHED/SUBMITTED

1. Wei Lin, Zhihua Qu, and Marwan A. Simaan, "*Nash strategies for pursuit-evasion differential games involving limited observation*", IEEE Trans. Aerospace and Electronic Systems, submitted.

2. Wei Lin, Zhihua Qu, and Marwan A. Simaan, "*Distributed ϵ-Nash Strategies for Multi-Pursuer Single-Evader Differential Games*", 2014 American Control Conference (ACC), submitted.

3. Ali Maknouninejad, Wei Lin, and Zhihua Qu, "*Optimum design and analysis of the cooperative control, applied to the distributed generators control in smart grids*" In IEEE PES Innovative Smart Grid Technologies (ISGT), pp. 1-6, 2013.

4. Wei Lin, Zhihua Qu, and Marwan A. Simaan, "*Multi-pursuer single-evader differential games with limited observations*", In American Control Conference (ACC), pp. 2711-2716, Washington D.C., 2013.

5. Ali Maknouninejad, Wei Lin, Hendra G. Harno, Zhihua Qu, and Marwan A. Simaan," *Cooperative control for self-organizing microgrids and game strategies for optimal dispatch of distributed renewable generations*", Energy Systems, vol. 3, no. 1, pp. 23-60, 2012.

6. Wei Lin, Zhihua Qu, and Marwan A. Simaan, "*A design of entrapment strategies for the distributed pursuit-evasion game*", In 18th IFAC World Congress, pp. 9334-9339, Milano, Italy. 2011.

7. Wei Lin, Zhihua Qu, and Marwan A. Simaan, "*A design of distributed nonzero-sum Nash strategies*", In 49th IEEE Conference on Decision and Control (CDC), pp. 6305-6310, Atlanta, Georgia, 2010.

LIST OF REFERENCES

[1] J. V. Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

[2] J. F. Nash *et al.*, "Equilibrium points in n-person games," *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[3] E. Bryson and Y. C. Ho, *Applied Optimal Control*. Bristol, PA, USA: Hemisphere Publishing Corp., 1975.

[4] L. Zadeh, "Optimality and non-scalar-valued performance criteria," *IEEE Trans. Automatic Control*, vol. 8, no. 1, pp. 59–60, 1963.

[5] R. Isaacs, *Differential Games*. John Wiley and Sons, 1965.

[6] L. Berkowitz and H. Fleming, "On differential games with integral payoff," *Ann. Math. Studies*, no. 39, pp. 413–435, 1957.

[7] A. Starr and Y. Ho, "Nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, pp. 184–206, 1969.

[8] E. Mageirou, "Values and strategies for infinite time linear quadratic games," *IEEE Trans. Automatic Control*, vol. 21, no. 4, pp. 547–550, 1976.

[9] D. Lukes, "Equilibrium feedback control in linear games with quadratic cost," *SIAM Journal on Control and Optimization*, vol. 9, pp. 234–252, 1971.

[10] T. Basar, "On the uniqueness of the nash solution in linear-quadratic differential games," *International Journal of Game Theory*, vol. 5, no. 2-3, pp. 65–90, 1976.

[11] G. Freiling, G. Jank, and H. Abou-Kandil, "On global existence of solutions to coupled matrix riccati equations in closed-loop nash games," *IEEE Trans. Automatic Control*, vol. 41, no. 2, pp. 264–269, 1996.

[12] M. Simaan and J. Cruz, "On the stackelberg solution in the nonzero-sum games," *Journal of Optimization Theory and Applications*, vol. 18, pp. 322–324, 1973.

[13] ——, "A stackelberg solution for games with many players," *Journal of Optimization Theory and Applications*, vol. 18, pp. 322–324, 1973.

[14] T. Basar, "On the relative leadership property of stackelberg strategies," *Journal of Optimization Theory and Applications*, vol. 11, pp. 655–661, 1973.

[15] J. Medanic, "Closed-loop stackelberg strategies in linear-quadratic problems," *IEEE Trans. Automatic Control*, vol. 23, no. 4, pp. 632–637, 1978.

[16] G. Papavassilopoulos and J. Cruz, "Nonclassical control problems and stackelberg games," *IEEE Trans. Automatic Control*, vol. 24, no. 2, pp. 155–166, 1979.

[17] G. Freiling, G. Jank, and S. Lee, "Existence and uniqueness of open-loop stackelberg equilibria in linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 110, no. 3, pp. 515–544, 2001.

[18] A. Weeren, J. Schumacher, and J. Engwerda, "Asymptotic analysis of linear feedback nash equilibria in nonzero-sum linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 101, no. 3, pp. 693–722, 1999.

[19] T. Basar and P. Bernhard, *H-infinity optimal control and related minimax design problems: a dynamic game approach*, 1st ed.   Birkhauser, 1995.

[20] E. Semsar-Kazerooni and K. Khorasani, "Multi-agent team cooperation: A game theory approach," *Automatica*, vol. 45, no. 10, pp. 2205–2213, 2009.

[21] D. Li and J. B. Cruz, "Defending an asset: a linear quadratic game approach," *IEEE Trans. Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1026–1044, 2011.

[22] K. G. Vamvoudakis, F. L. Lewis, and G. R. Hudas, "Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality," *Automatica*, 2012.

[23] A. Maknouninejad, W. Lin, H. Harno, Z. Qu, and M. Simaan, "Cooperative control for self-organizing microgrids and game strategies for optimal dispatch of distributed renewable generations," *Energy Systems*, vol. 3, no. 1, pp. 23–60, 2012.

[24] Y. Ho, A. Bryson, and S. Baron, "Differential games and optimal pursuit-evasion strategies," *IEEE Trans. Automatic Control*, vol. AC-10, no. 4, pp. 385–389, 1965.

[25] M. Foley and W. Schmitendorf, "A class of differential games with two pursuers versus one evader," *IEEE Trans. Automatic Control*, vol. 19, pp. 239–243, 1974.

[26] Z. E. Fuchs, P. P. Khargonekar, and J. Evers, "Cooperative defense within a single-pursuer, two-evader pursuit evasion differential game," in *49th IEEE Conference on Decision and Control (CDC)*, Atlanta, GA, 2010, pp. 3091–3097.

[27] D. Li and J. Cruz, "Improvement with look-ahead on cooperative pursuit games," in *45th IEEE Conference on Decision and Control*, San Diego, CA, 2006, pp. 4313–4318.

[28] X. Wang, J. B. Cruz Jr, G. Chen, K. Pham, and E. Blasch, "Formation control in multi-player pursuit evasion game with superior evaders," in *Defense and Security Symposium*. International Society for Optics and Photonics, 2007, pp. 657 811–657 811.

[29] J. Guo, G. Yan, and Z. Lin, "Local control strategy for moving-target-enclosing under dynamically changing network topology," *Systems & Control Letters*, vol. 59, pp. 654–661, 2010.

[30] W. Lin, Z. Qu, and M. Simaan, "A design of entrapment strategies for the distributed pursuit-evasion game," in *18th IFAC World Congress*, Milan, Italy, 2011, pp. 9334–9339.

[31] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "A cooperative homicidal chauffeur game," *Automatica*, vol. 45, no. 7, pp. 1771–1777, 2009.

[32] J. S. Jang and C. J. Tomlin, "Control strategies in multi-player pursuit and evasion game," *AIAA*, vol. 6239, pp. 15–18, 2005.

[33] H. Huang, W. Zhang, J. Ding, D. M. Stipanovic, and C. J. Tomlin, "Guaranteed decentralized pursuit-evasion in the plane with multiple pursuers," in *50th IEEE Conference on Decision and Control and European Control Conference*.   IEEE, 2011, pp. 4835–4840.

[34] S. Bhattacharya and T. Başar, "Differential game-theoretic approach to a spatial jamming problem," in *Advances in Dynamic Games*.   Springer, 2013, pp. 245–268.

[35] X. Wang, J. Cruz, G. Chen, K. Pham, and E. Blasch, "Formation control in multi-player pursuit evasion game with superior evaders," in *Defense Transformation and Net-Centric Systems*, vol. 6578, Orlando, Florida, USA, 2007, pp. 657 811.1–657 811.9.

[36] M. Wei, G. Chen, J. B. Cruz, L. Haynes, K. Pham, and E. Blasch, "Multi-pursuer multi-evader pursuit-evasion games with jamming confrontation," *Journal of Aerospace Computing, Information, and Communication*, vol. 4, no. 3, pp. 693–706, 2007.

[37] S. Bopardikar, F.Bullo, and J. Hespanha, "On discrete-time pursuit-evasion games with sensing limitations," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1429–1439, 2008.

[38] A. Antoniades, H. J. Kim, and S. Sastry, "Pursuit-evasion strategies for teams of multiple agents with incomplete information," in *42nd IEEE Conference on Decision and Control*, vol. 1, 2003, pp. 756–761.

[39] D. Li and J. Cruz, "Graph-based strategies for multi-player pursuit evasion games," in *46th IEEE Conference on Decision and Control*. IEEE, 2007, pp. 4063–4068.

[40] W. Lin, Z. Qu, and M. Simaan, "A design of distributed nonzero-sum nash strategies," in *49th IEEE Conferenceon Decision and Control*, Atlanta, USA, 2010, pp. 6305–6310.

[41] W. Lin, Z. Qu, and M. A. Simaan, "Multi-pursuer single-evader differential games with limited observations," in *American Control Conference*, Washington D.C., USA, 2013, pp. 2711–2716.

[42] P. Cheng, "A short survey on pursuit-evasion games," Department of Computer Science, University of Illinois at Urbana-Champaign, Tech. Rep., 2003.

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

[44] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control: Collective group behavior through local interaction," *IEEE Control Syst. Mag.*, vol. 27, pp. 71–82, 2007.

[45] H. Kim, H. Shim, and J. H. Seo, "Output consensus of heterogeneous uncertain linear multi-agent systems," *IEEE Trans. Automatic Control*, vol. 56, no. 1, pp. 200 –206, jan. 2011.

[46] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. London: Springer Verlag, 2009.

[47] D. M. Stipanovic, G. Inalhan, R. Teo, and C. J. Tomlin, "Decentralized overlapping control of a formation of unmanned aerial vehicles," *Automatica*, vol. 40, no. 8, pp. 1285 – 1296, 2004.

[48] T. Keviczky, F. Borrelli, K. Fregene, D. Godbole, and G. Balas, "Decentralized receding horizon control and coordination of autonomous vehicle formations," *IEEE Trans. Control System Technology*, vol. 16, no. 1, pp. 19–33, 2008.

[49] J. Wang and M. Xin, "Integrated optimal formation control of multiple unmanned aerial vehicles," *IEEE Trans. Control System Technology*, vol. PP, no. 99, pp. 1–1, 2012.

[50] Z. Li, Z. Duan, G. Chen, and L. Huang, "Consensus of multiagent systems and synchronization of complex networks: a unified viewpoint," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 57, no. 1, pp. 213–224, 2010.

[51] C. Li and Z. Qu, "Cooperative attitude synchronization for rigid-body spacecraft via varying communication topology," *International Journal of Robotics and Automation*, vol. 26, no. 1, p. 110, 2011.

[52] P. Wieland, R. Sepulchre, and F. Allg??wer, "An internal model principle is necessary and sufficient for linear output synchronization," *Automatica*, vol. 47, no. 5, pp. 1068 – 1074, 2011.

[53] H. F. Grip, T. Yang, A. Saberi, and A. A. Stoorvogel, "Output synchronization for heterogeneous networks of non-introspective agents," *Automatica*, vol. 48, no. 10, pp. 2444 – 2453, 2012.

[54] W. Levine and M. Athans, "On the determination of the optimal constant output feedback gains for linear multivariable systems," *IEEE Trans. Automatic Control*, vol. 15, no. 1, pp. 44–48, 1970.

[55] W. Levine, T. Johnson, and M. Athans, "Optimal limited state variable feedback controllers for linear systems," *IEEE Trans. Automatic Control*, vol. 16, no. 6, pp. 785–793, 1971.

[56] V. Blondel and J. Tsitsiklis, "Np-hardness of some linear control design problems," *SIAM Journal on Control and Optimization*, vol. 35, no. 6, pp. 2118–2127, 1997.

[57] V. Syrmos, C. Abdallah, P. Dorato, and K. Grigoriadis, "Static output feedback-a survey," *Automatica*, vol. 33, no. 2, pp. 125–137, 1997.

[58] F. Lewis, D. Vrabie, and V. Syrmos, *Optimal Control*, 3rd ed.   Wiley-Interscience, 2012.

[59] K. Chu, "Optimal decentralized regulation for a string of coupled systems," *IEEE Trans. Automatic Control*, vol. 19, no. 3, pp. 243–246, 1974.

[60] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Trans. Automatic Control*, vol. 47, no. 7, pp. 1091–1107, 2002.

[61] M. Jovanovic, "On the optimality of localized distributed controllers," *Int. J. Systems, Control and Communications*, vol. 2, no. 1/2/3, pp. 82–99, 2010.

[62] X. Qi, M. Salapaka, P. Voulgaris, and M. Khammash, "Structured optimal and robust control with multiple criteria: A convex solution," *IEEE Trans. Automatic Control*, vol. 49, no. 10, pp. 1623–1640, 2004.

[63] M. Rotkowitz and S. Lall, "A characterization of convex problem in decentralized control," *IEEE Trans. Automatic Control*, vol. 51, no. 2, pp. 273–286, 2006.

[64] V. Gupta, B. Hassibi, and R. Murray, "A sub-optimal algorithm to synthesize control laws for a network of dynamic agents," *International Journal of Control*, vol. 78, no. 16, pp. 1302–1313, 2005.

[65] F. Borrelli and T. Keviczky, "Distributed lqr design for identical dynamically decoupled systems," *IEEE Trans. Automatic Control*, vol. 53, no. 8, pp. 1901–1912, 2008.

[66] D. Gu, "A differential game approach to formation control," *IEEE Trans. Control Systems Technology*, vol. 16, no. 1, pp. 85–93, 2008.

[67] E. Semsar-Kazerooni and K. Khorasani, "Multi-agent team cooperation: A game theory approach," *Automatica*, vol. 45, no. 10, pp. 2205 – 2213, 2009.

[68] D. Gu, "A game theory approach to target tracking in sensor networks," *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 1, pp. 2–13, 2011.

[69] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Trans. Industrial Informatics*, vol. 9, no. 1, pp. 427–438, 2013.

[70] F. R. Chung, *Spectral graph theory*. AMS Bookstore, 1997, vol. 92.

[71] L. S. Pontryagin, *The mathematical theory of optimal processes*. CRC Press, 1962, vol. 4.

[72] S. J. Qin and T. A. Badgwell, "An overview of industrial model predictive control technology," in *AIChE Symposium Series*, vol. 93, no. 316. New York, NY: American Institute of Chemical Engineers, 1971-c2002., 1997, pp. 232–256.

[73] M. Simaan and J. Cruz Jr, "Sampled-data nash controls in non-zero-sum differential games," *International Journal of Control*, vol. 17, no. 6, pp. 1201–1209, 1973.

[74] E. K. Chong and S. H. Zak, *An introduction to optimization*. John Wiley & Sons, 2013, vol. 76.

[75] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in MATLAB," in *CACSD Conference*, Taipei, Taiwan, 2004.

[76] C. Crusius and A. Trofino, "Sufficient lmi conditions for output feedback control problems," *IEEE Trans. Automatic Control*, vol. 44, no. 5, pp. 1053–1057, 1999.

[77] P. Menon, G. Sweriduk, and B. Sridhar, "Optimal strategies for free-flight air traffic conflict resolution," *Journal of Guidance, Control, and Dynamics*, vol. 22, no. 2, pp. 202–211, 1999.

[78] Y. Xu, "Nonlinear robust stochastic control for unmanned aerial vehicles," *Journal of guidance, control, and dynamics*, vol. 32, no. 4, pp. 1308–1319, 2009.

[79] "Advisory circular," U.S. Department of Transportation and Federal Aviation Administration, Tech. Rep. AC 120-28D, 1999.