

APPROXIMATE FEEDBACK SOLUTIONS FOR DIFFERENTIAL GAMES

THEORY AND APPLICATIONS

by
THULASI MYLVAGANAM

A Thesis submitted in fulfillment of requirements for the degree of
Doctor of Philosophy

Control and Power Research Group
Department of Electrical and Electronic Engineering
Imperial College London
2014

Abstract

Differential games deal with problems involving multiple players, possibly competing, that influence common dynamics via their actions, commonly referred to as strategies. Thus, differential games introduce the notion of strategic decision making and have a wide range of applications.

The work presented in this thesis has two aims. First, constructive approximate solutions to differential games are provided. Different areas of application for the theory are then suggested through a series of examples. Notably, multi-agent systems are identified as a possible application domain for differential game theory. Problems involving multi-agent systems may be formulated as nonlinear differential games for which closed-form solutions do not exist in general, and in these cases the constructive approximate solutions may be useful.

The thesis is commenced with an introduction to differential games, focusing on feedback Nash equilibrium solutions. Obtaining such solutions involves solving coupled partial differential equations. Since closed-form solutions for these cannot, in general, be found two methods of constructing approximate solutions for a class of nonlinear, nonzero-sum differential games are developed and applied to some illustrative examples, including the multi-agent collision avoidance problem. The results are extended to a class of nonlinear Stackelberg differential games. The problem of monitoring a region using a team of agents is then formulated as a differential game for which ad-hoc solutions, using ideas introduced previously, are found. Finally mean-field games, which consider differential games with infinitely many players, are considered. It is shown that for a class of mean-field games, solutions rely on a set of ordinary differential equations in place of two coupled partial differential equations which normally characterise the problem.

Declaration of Originality

I hereby declare that the work presented in this thesis is my own unless otherwise stated. To the best of my knowledge the work is original and ideas developed in collaboration with others have been appropriately referenced.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgments

This thesis is the result of research work performed at the department of electrical and electronic engineering, Imperial College London. This work would not have been possible without the collaboration and support of several people to whom I would like to express my gratitude.

First, I would like to thank my supervisor, Prof. Astolfi for his support, guidance and help. I could not have asked for a better supervisor, and consider myself fortunate and am grateful for having had the opportunity to work with him. I have learnt a lot from him over the last years and, moreover, I have enjoyed myself while doing so.

I would like to thank my examiners Prof. De Persis and Prof. Parisini for their constructive feedback.

My PhD has been partly funded by the Austrian Centre for Competance in Mechatronics and during the last three years, I spent some time at the Johannes Kepler University, Linz. I would like to thank Prof. Del Re and Thomas E. Passenbrunner for their collaboration and I would also like to thank them and Daniela Hummer for making me feel welcome during the time I spent in Linz.

During the last year, I have had the opportunity to work with Prof. Bauso, who introduced me to the topic of mean-field games. I would like to thank him for his collaboration and for sharing his passion for the topic.

Over the last three years I have had the chance to work with my friend and colleague Mario Sassano. I would like to thank him for his advice and good teamwork. Working with Mario and Alessandro has been a real pleasure and I am lucky for having had the experience of doing so.

Furthermore, I have been fortunate to meet many friends, both within and outside the research group, who helped create a pleasant working and living environment.

The last years would not have been as enjoyable without them. For this I would like to thank Sebastian Krumscheid, Mario Sassano, Giordano Scarciotti, Jelena Dragovic, Claudia Spallarossa, Caitríona Sheridan, Nastasia Bassil, Claudia Battistelli, Lauriane Thorner, Andrea Boccia and the rest of the control and power research group.

I would like to thank my sister Janani, Mari Berg and Silje Håvardstun for their friendship. They are always ready to lend an ear when it is necessary or remind me to laugh if I happen to be too serious, and this has helped me through many challenges.

Finally, I would like to thank my parents Haima and Saba, and Janani for their unconditional love and support. Their constant encouragement and support is invaluable to me and gives me happiness and peace of mind. I would not have been able to achieve these results without them.

I would like to express my deepest gratitude to all these amazing people for having made the last years unforgettable.

To my parents.

Contents

Declaration of Originality	5
Copyright Declaration	7
Acknowledgments	9
Contents	13
List of Figures	17
Chapter 1. Introduction	21
1.1 Motivation and Objectives	21
1.2 Contributions	28
1.3 Notation	30
1.4 Published Results	31
Chapter 2. Introduction and Preliminaries to Differential Game Theory	33
2.1 Introduction	33
2.2 Problem Formulation	36
2.2.1 2-Player Differential Games	36
2.2.2 N -player differential games	40
2.3 Linear-Quadratic Differential Games	42
2.3.1 2-player Linear-Quadratic Differential Games	44
2.3.2 N -player Linear-Quadratic Differential Games	45
2.3.3 Some Remarks Concerning Linear-Quadratic Differential Games .	47
2.4 ϵ -Nash Equilibria	48
2.5 Conclusion	49
Chapter 3. Constructive ϵ-Nash Equilibrium Solutions for Nonzero-Sum Differential Games	51
3.1 Introduction	51
3.2 Nash and ϵ -Nash Equilibria for Non-cooperative differential games	53

3.2.1	2-Player Problem Formulation	54
3.2.2	N -Player Problem Formulation	58
3.3	Linear-Quadratic Non-Cooperative Differential Games	62
3.4	Algebraic $\bar{P}$ Matrix Solution	65
3.4.1	2-Player Algebraic $\bar{P}$ Matrix Solution	65
3.4.2	N -Player Algebraic $\bar{P}$ Matrix Solution	66
3.5	Approximate Solutions Using Individual Dynamic Extensions	67
3.5.1	2-Player Case: Approximate Solutions Using Two Individual Dy- namic Extensions	68
3.5.2	N -Player Case: Approximate Solutions Using N Individual Dy- namic Extensions	74
3.6	Approximate Solutions Using a Shared Dynamic Extensions	77
3.6.1	2-Player Case: Approximate Solutions Using a Shared Dynamic Ex- tension	77
3.6.2	N -Player Case: Approximate Solutions Using a Shared Dynamic Extension	82
3.7	Numerical Examples	84
3.7.1	Academic Example	84
3.7.2	Application to Lotka-Volterra Models	90
3.8	Applications to Multi-Agent Systems	95
3.8.1	Problem Formulation	96
3.8.2	Solving the Multi-Agent Collision Avoidance Problem	99
3.8.3	Multi-Agent Collision Avoidance	101
3.8.4	Simulations	104
3.9	Conclusion	111
Chapter 4. 2-Player Stackelberg Differential Games		113
4.1	Introduction	113
4.2	The Two-Player Stackelberg Differential Game	115
4.3	Algebraic $\bar{P}$ Matrix Stackelberg Solution	120
4.4	Approximate Solutions Using a Shared Dynamic Extension	121
4.5	Simulations	127
4.6	Conclusion	131
Chapter 5. Optimal Monitoring		133
5.1	Introduction	133
5.2	Multi-Agent Systems and Monitoring Tasks	135
5.3	Multi-Agent problems Formulated as Differential Games	137

5.4	Problem Formulation	139
5.5	Special case: Single Agent on a Segment	145
5.6	Approximate Solutions to the Optimal Control Problems	147
5.7	Unicycle Agent Dynamics	151
5.8	Simulations	154
5.8.1	Approximate Solution to One Optimal Control Problem	155
5.8.2	Monitoring by Solving a Sequence of Optimal Control Problems	156
5.9	Conclusion	160
Chapter 6.	Mean-Field Games	161
6.1	Introduction	161
6.2	Background	163
6.3	Mean-Field Games and Two-Point Boundary Value Problems	166
6.3.1	Problem Formulation	167
6.3.2	Main results	169
6.3.3	Interpretation of Results	175
6.4	Simulations	177
6.4.1	Quadratic Density Distribution	177
6.4.2	Fourth Order Polynomial Distribution	178
6.5	Conclusion	180
Chapter 7.	Conclusion	185
7.1	Summary of contributions	185
7.2	Future	186
Bibliography		189

List of Figures

3.1	Comparison of state trajectories using different strategies.	88
3.2	Comparison of state trajectories using different strategies.	88
3.3	Comparison of state trajectories using different strategies.	88
3.4	Comparison of state trajectories using different strategies.	88
3.5	Comparison of state trajectories using different strategies.	89
3.6	Regions in which each strategy is <i>closer</i> to the Nash equilibrium strategies for player 1.	90
3.7	Regions in which each strategy is <i>closer</i> to the Nash equilibrium strategies for player 2.	90
3.8	Time histories of dynamic extensions used for the dynamic feedback laws.	91
3.9	Comparison of linear strategies and dynamic strategies using several dy- namic extensions.	94
3.10	Comparison of linear strategies and dynamic strategies using a shared dy- namic extension.	94
3.11	Comparison of linear strategies and dynamic strategies using several dy- namic extensions.	94
3.12	Comparison of linear strategies and dynamic strategies using a shared dy- namic extension.	94
3.13	Comparison of state trajectory for the different strategies.	95
3.14	Trajectories of two agents swapping positions while avoiding collisions. .	105
3.15	Time history of distance between two agents swapping positions.	106
3.16	Time histories of agents' control strategies.	107
3.17	Trajectories of two agents avoiding collisions.	108
3.18	Time history of distance between two agents avoiding collisions.	108
3.19	Time histories of agents' control strategies.	109
3.20	Trajectories of three agents avoiding collisions.	109
3.21	Time histories of distance between three agents avoiding collisions.	110
3.22	Time histories of agents' control strategies.	110

4.1	Comparison of strategies in a Stackelberg differential game for player 1.	130
4.2	Comparison of strategies in a Stackelberg differential game for player 2.	130
4.3	Comparison of state trajectories using different strategies in a Stackelberg differential game.	131
5.1	Possible sensor models for the <i>optimal monitoring problem</i>	141
5.2	Time history of the position of an agent monitoring for the 1D coverage problem.	147
5.3	The values of the virtual player and the coverage map for the 1D coverage problem.	148
5.4	Time histories of the minimum and maximum coverage levels across the 1D region.	148
5.5	Trajectories of two agents resulting from the (approximate) solution of <i>one</i> optimal control problem.	155
5.6	Time histories of errors between the <i>trajectory plan</i> and trajectories taken by the agents with unicycle dynamics.	156
5.7	Time histories of the distance between the agents (for agents with single-integrator dynamics and unicycle dynamics).	157
5.8	Trajectories of two agents monitoring a region.	158
5.9	Time histories of error between <i>trajectory plan</i> and trajectories taken by the agents with unicycle dynamics.	158
5.10	Coverage map at the end of the mission.	159
5.11	Time histories of the maximum and minimum coverage levels.	159
6.1	Evolution of quadratic density distributions for converging and diverging vector fields.	177
6.2	Time histories of coefficients of quadratic density distribution and value function.	179
6.3	Density distribution of agents (quadratic case) without noise.	180
6.4	Density distribution of agents (quadratic case) with noise.	181
6.5	The first three coefficients of (fourth order) polynomial density distribution and quadratic value functions.	182
6.6	The last two coefficients of (fourth order) polynomial density distribution.	182
6.7	Density distribution of agents without noise.	183
6.8	Density distribution of agents with some noise.	183
6.9	Density distribution of agents with significant noise.	184

6.10 Time histories of the state of each player for different noise levels.	184
---	-----

Chapter 1

Introduction

1.1 Motivation and Objectives

Control theory is the study of dynamical systems and their behaviours. Given a mathematical model of such a system the idea is, as the name suggests, to *control* the system in some way. It may, for example, be of interest to render an equilibrium stable or to ensure that the output tracks a desired reference signal. Control theory provides tools to identify if a certain behaviour is achievable, and if so how it can be achieved. Because of its abstract nature, control theory is interdisciplinary and thus plays a fundamental role in many fields of engineering and mathematics.

For example, control theory is important in fields such as robotics, automation and electrical engineering. It plays an important role in power generation and in maintaining the safe operation of power plants. In nuclear power plants, for instance, it is common to use remote-controlled technologies to handle radioactive waste, thus minimising the exposure of human beings to toxic materials. Furthermore, in the event of an accident, control systems are often central in managing damage. For example, after the recent nuclear disaster in Fukushima robots have been used to monitor and clean the areas affected by the nuclear leak [1]. With the advance of robotics and small-scale electronic devices, control theory plays a growing role in biomedical engineering: robotic assisted surgery is becoming more common, as are robotic prosthetics [2–4]. These “traditional” applica-

tions aside, control theory appears in many different areas too. For instance, in [5] the authors apply control theoretic methods to enhance the immune system of patients infected with the human immunodeficiency virus (HIV) infection. More precisely, a mathematical model of the development of the HIV infection in a patient shows that there are two asymptotically stable equilibria, only one of which is favourable. In [5] drug therapy scheduling in order to drive the patient to the favourable equilibrium is proposed using a reduced-order model of the HIV infection dynamics. It is evident that control theory has a wide range of applications, many of which are of great importance to society.

On another front, since the 1940's mathematicians, with Von Neumann and Morgenstern at the forefront, have studied the field of game theory. Game theory is the study of problems involving several players and the decisions these players make to optimise their own performance criteria [6, 7]. In the well-known zero-sum games, these performance criteria are purely competitive in the sense that a gain for one player must mean an equivalent loss for other players. A large variety of "typical" games, such as poker and standard pursuit-evasion games, can be described by this class of games. For a 2-player zero-sum game, a win for one player implies that the other player inevitably loses the game which, for example, is the case in a game of chess. Nonzero-sum games, on the other hand, generalise this concept by allowing the players to have individual goals that may, or may not be, conflicting. Therefore, in this framework a gain for one player need not imply a loss for any of the other players. In both nonzero- and zero-sum games, given a group of players with their own goals, game theory deals with determining the "best" strategies for each of the players. *Zur Theorie der Gesellschaftsspiele*, published in 1928 by Von Neymann and Morgenstern, is often considered the first paper on game theory and *The Theory of Games and Economic Behaviour* published in 1944, by the same authors, is pivotal to the formalisation of game theory. However, ideas relating to the topic also appeared previously, notably by Cournot as early as in 1838 and by Stackelberg in 1935 [8, 9]. Different solution concepts for a game exist, the most common one being the so-called Nash equilibrium solution, named after John Nash, who won a Nobel prize for his work on game theory in 1994 [10–12].

Whereas game theory deals with static problems, differential games are their dynamic counterpart. That is, differential games are concerned with situations in which several players influence a common dynamical system, via their *control strategies*, to attain some predefined goals. Notions of solutions and equilibria are identical to those in game theory, however additional care must be taken to ensure stability properties. In general a differential game is characterised by a dynamical system and a set of N players each associated with a cost functional that *quantifies* the goal of the player.

Since both control theory and differential games deal with dynamical systems one might wonder whether the two fields have more in common and whether the framework provided by differential game theory can be of use to control theory. To answer this question consider first the well-established topic of optimal control. In optimal control, given a dynamical system, the problem of determining the best control input to minimise (or maximise) a cost functional is considered, *i.e.* the problem is to determine the best strategy of a single decision maker. Differential games, on the other hand, can be viewed as an extension of this in which, instead of one decision maker, there are several decision makers. These decision makers are referred to as players, each seeking to minimise (or maximise) their own cost functionals, subject to the dynamics of the system. Thus, differential games introduce the notion of *strategic behaviour* to control theory. This is naturally useful for control problems in which a dynamical system has several inputs that may have different goals, such as what is seen in robust control. H_∞ control considers problems in which a dynamical system is influenced by an adversarial disturbance in addition to the control input. The control input is then designed to guarantee certain performance criteria, which is often done by formulating the robust control problem as a zero-sum differential game, in which the control input and the adversarial disturbance are “opposing players” [13]. Similarly, mixed H_2/H_∞ control is concerned with selecting a control input guaranteeing given performance criteria, while simultaneously attempting to optimise a given cost functional. Such problems can be formulated as nonzero-sum differential games as done on [14] for linear systems and [15] for nonlinear systems. However, differential games can be useful to other areas of control theory as well. Some notable examples, which illustrate possible areas of applications of differential game theory, are

introduced and studied in this thesis. To motivate the theory developed in the remainder of the thesis, a brief introduction and description of some of these examples are provided here.

The first example to make a point of at this stage is the differential game involving a so-called Lotka-Volterra model considered in Section 3.7.2. In this example a biological system consisting of two competing species is considered. The population of the two species are represented by the states, x_1 and x_2 , of the system and their dynamics are given by

$$\dot{x}_1 = b_1x_1 - a_{11}x_1^2 - a_{21}x_1x_2,$$

$$\dot{x}_2 = b_2x_2 - a_{12}x_1x_2 - a_{22}x_2^2,$$

where $b_i > 0$ and $a_{ij} > 0$, for $i = 1, 2$ and $j = 1, 2$, are birth rates and predation efficiencies, respectively. We consider the case in which there are two players which can influence the predation efficiencies via control inputs, u_{ij} , where $i = 1, 2$ and $j = 1, 2$, in such a way that the efficiencies are given by $a_{ij} = a_{ij}^* + u_{ij}$, where $a_{ij}^* > 0$ with $i = 1, 2$ and $j = 1, 2$. The input $u_i = (u_{i1}, u_{i2})^\top$ is the control strategy of player i and we assume that that players 1 seeks to drive the population x_2 to a value $x_2^* > 0$ whereas player 2 seeks to drive the population x_1 to a value $x_1^* > 0$. Furthermore, each player attempts to do so while minimising its own effort and attempting to maximise the efforts of the other player. This problem describes a competitive biological system and illustrates the possible application of differential game theory to analyse and synthesise the behaviour of certain biological systems.

The *multi-agent collision avoidance problem*, introduced in Section 3.8, describes the task of maneuvering a team of N agents with single-integrator dynamics from given initial positions, $x_i(0) \in \mathbb{R}^2$, $i = 1, \dots, N$, to predefined target positions, denoted by $x_i^* \in \mathbb{R}^2$, $i = 1, \dots, N$, while avoiding inter-agent collisions. Furthermore, each agent attempts minimising its efforts while doing so. The problem is formulated as a nonzero-sum differential game where agent i seeks to minimise a cost functional of the form

$$J_i(u_1, \dots, u_N) = \frac{1}{2} \int_0^\infty \left((\alpha_i + \beta_i g_i(x_1, \dots, x_N))(x_i - x_i^*)^\top (x_i - x_i^*) + \|u_i(t)\|^2 \right) dt,$$

where $x_i(t) \in \mathbb{R}^2$ is the position of agent i at time $t > 0$, $\alpha_i > 0$ and $\beta_i > 0$ are constants and the function $g_i(x_1, \dots, x_N)$, which is discussed in more detail in Section 3.8, is such that it *encourages* the agent to maintain a so-called *safety distance* between itself and other agents. Approximate solutions for the resulting nonzero-sum differential game, in terms of Nash equilibrium strategies, are found using the theory presented in Section 3.6. The *multi-agent collision avoidance problem* demonstrates how a problem involving several agents can be formulated as a differential game.

As the name suggests, multi-agent systems are systems consisting of several *agents*, which may in general be heterogeneous. The agents may be any dynamical systems, such as, for example, mobile sensors, robots, unmanned vehicles (UVs) or individuals in a social network [16]. Multi-agent systems have gained a significant level of attention in recent years which may be attributed to several factors, two of which may be considered as the main motivators [16]. First, many systems occurring in nature, for example in biology or material sciences, consist of individual elements that have a collective functionality, and the behaviour of such complex systems may be understood by considering them as networked multi-agent systems [16–19]. Secondly, the ever increasing availability of cheap autonomous robots has resulted in an interest in the synthesis of networked engineering systems, such as multi-vehicle systems and sensor networks [16]. In fact, many engineering applications of multi-agent systems take inspiration from biological systems: in [20] the behaviour of a school of fish is used as an inspiration to develop a coordinated control strategy for distributed sensors to sample an environment. A similar problem is studied in [21], where the coverage of a planar region by a multi-agent system, which could represent a sensor network or animal groups foraging for food, is studied. Multi-agent systems are also used to implement intelligent swarm behaviour, often drawing inspiration from the motion of swarms of birds [22, 23]. Other areas of application include military and defense, power systems, sensor networks and search and rescue missions, social networks, quantum networks and nanostructures [16, 19].

System of systems is another area that has gained interest within the control community in recent years and plays a significant role in many applications, such as military,

security, aerospace, manufacturing and environmental systems [24,25]. Different definitions of a system of systems exist, but in essence this is a collection of individual systems that are integrated together to solve a task collaboratively. The individual systems may be either heterogeneous or homogeneous and the idea behind integrating them together may, for example, be to enhance certain aspects, such as robustness or reliability, of the overall system. A system of systems approach provides a high-level point-of-view to study problems involving a group of interconnected, individual systems. A multi-agent system can be thought of as a particular example of a system of systems, where each agent is an individual system, whereas the team of agents can be thought of as a *overall* system, yielding the term *system of systems*. In [26,27], for example, a swarm of cooperating robots is considered as a system of systems. Whichever perspective one may choose to study multi-agent systems, many open problems exist. For example more research is needed to extend standard notions from traditional control theory, such as stability, controllability, observability, etc., to the high-level system of systems [24]. In a system consisting of several agents, local and global stability are not the same and neither is individual optimality and group optimality [19]. Thus, the study of local interactions and collective behaviour introduce challenges both in terms of analysis and synthesis of networked multi-agent systems [16,19]. In fact, in [24], the authors even suggest that the research on complex systems may be the birth of a new engineering discipline which they term “Systems-of-systems engineering”.

Systems consisting of several *agents* can perform complex tasks that would otherwise be difficult, dangerous or even impossible. For example, in natural systems collective behaviour, such as flocking, can protect individuals in a group from predators and it is well-known that migrating birds fly in a so-called “V-formation” to minimise the effort required by each individual bird [19]. In engineered systems it is typically desirable for the team of agents to collectively perform a task. Autonomous multi-agent systems can perform useful tasks without human involvement and as a result treacherous tasks can be undertaken without directly involving and endangering human life. For example, in the event of a natural disaster a team of autonomous agents with sensors can gather information that could aid search and rescue missions without directly involving or endangering

the lives of a rescue team. This is what motivates the work in Chapter 5, where the problem of continuously monitoring a region using a team of agents equipped with sensors is formulated as a differential game, for which ad-hoc solutions are obtained. Both this problem and the multi-agent collision avoidance problem suggests multi-agent systems as a possible application domain for the framework provided by differential game theory. In fact, since multi-agent systems consist of several, possibly heterogeneous, agents that solve tasks, either cooperatively or competitively, simply interchanging the words “agents” and “players”, it appears natural that problems involving multi-agent control can be formulated as differential games and, more precisely, as nonzero-sum differential games.

In some applications a differential game may have many participating players, or a multi-agent system may consist of a large number of agents. It will become apparent that as the number of players participating in a differential game grows large, obtaining solutions becomes increasingly cumbersome. Drawing inspiration from statistical physics, *mean-field games* consider differential games with an infinite number of indistinguishable players. Mean-field games were introduced independently by J.M. Lasry and P.L. Lions in [28] and by M.Y. Huang, P.E. Caines and R.P. Malhamé in [29,30]. In some cases it is more feasible to obtain solutions to the mean-field counterpart of a N -player differential game than the differential game itself. Thus, solutions to mean-field games can be used to approximate the solutions of N -player differential games for a class of games in which the N players are “small” and identical. In the limit as N tends to infinity the approximate solutions become exact. Mean-field games can, for example, be of use when considering problems involving social networks or distributed power systems with renewable energy sources.

The aim of this thesis is twofold. First, since it is generally not possible to obtain closed-form solutions to differential games, methods for constructing approximate solutions to differential games are developed. Second, the use of differential game theory to formulate and solve problems is demonstrated on the aforementioned notable examples, thus illustrating possible areas of applications for the theory.

1.2 Contributions

First, differential games are considered, starting with a brief background on differential games is provided in Chapter 2. Here general 2- and N -player differential games and their solutions, in terms of *feedback Nash equilibria*, are formally defined. The Nash equilibrium is the most common solution concept associated with differential games. Assuming all players are rational, a set of strategies is said to be a Nash equilibrium solution if it is such that a player cannot achieve a better performance by deviating from its Nash equilibrium strategy. This solution concept assumes that all players have full information regarding the state of the system, is aware of the cost functionals all of the players seek to minimise and that the players announce their strategies simultaneously. A special class of differential games, namely *linear-quadratic differential games* are then considered in Section 2.3. Obtaining equilibrium strategies for linear-quadratic differential games rely on the solution of coupled algebraic Riccati equations. However, even in this case obtaining solutions is not generally straight-forward, which motivates the construction of approximate solutions for differential games. The aim of this chapter is to provide the background necessary to follow the remainder of the material in Chapters 3 and 4. For a more comprehensive introduction to differential games the reader is referred to [6,31–33].

In Chapter 3 a class of nonlinear nonzero-sum differential games are considered. Once again, the differential games and their feedback Nash equilibrium solutions are formally defined. The standard notion of α -admissible strategies is defined before we introduce the somewhat different notion of ϵ_α -Nash equilibrium solutions. Some properties relating to linear-quadratic differential games are then brought to light. In particular an ϵ_α -Nash equilibrium solution is a set of strategies, $u^* = \{u_1^*, \dots, u_N^*\}$, which ensures that if any one player deviates from its ϵ_α -Nash equilibrium strategy and the resulting set of strategies is such that the zero equilibrium of the closed-loop system is locally exponentially stable and that the decay rate can be bounded from above by a constant $\alpha > 0$, the player can gain no more than $\epsilon_\alpha > 0$ compared to its outcome corresponding to the set u^* . The possible additional gain ϵ_α is parametrised with respect to the minimum decay rate α , and if $\epsilon_\alpha = 0$, the ϵ_α -Nash equilibrium solution coincides with the Nash

equilibrium solution. We show that for linear-quadratic differential games solving coupled algebraic Riccati inequalities allows us to determine ϵ_α -Nash equilibrium solutions where the value ϵ_α can be found by solving a Lyapunov equation as detailed in Section 3.3. The notion of *algebraic $\bar{P}$ matrix solution* is then defined in Section 3.4. This provides a tool that allows for the construction of dynamic feedback laws that constitute local ϵ_α -Nash equilibrium solutions for a differential game, as discussed in Section 3.5 and 3.6, *i.e.* in these sections two methods for constructing approximate solutions to nonlinear differential games are developed. In this case the ϵ_α term is the upper bound of an integral, which may not, in general, be straight-forward to evaluate. In Section 3.7 two numerical examples are considered and the dynamic approximate solutions are compared to the linear-quadratic approximations of the differential games. As a final example one of the developed methods is applied to a problem involving a multi-agent system in Section 3.8, where to so-called *multi-agent collision avoidance problem* is introduced. The multi-agent collision avoidance problem is defined as a differential game, an *algebraic $\bar{P}$ matrix solution* for the game is identified and used to construct local ϵ_α -Nash equilibrium strategies and the theory is illustrated by a set of simulations. As mentioned previously, this example suggests that multi-agent systems as a possible application domain for the theory.

In Chapter 4, a class of 2-player Stackelberg differential games, *i.e.* 2-player differential games with a certain *hierarchy* between the players, is considered. The methods of constructing dynamic feedback strategies constituting ϵ_α -Nash equilibrium strategies for Nash differential games are extended to this class of hierarchical differential games. The information structure associated with this solution concept is inherently different from that associated with Nash equilibrium solutions. When seeking Stackelberg equilibrium solutions it is assumed that the players announce their strategies sequentially and it is not assumed that all players know the performance criteria of all other players [11]. This solution concept may, for example, be useful when formulating problems in which some players are able to announce their strategies faster than the other players, possible due to differences in computational power available for the players. The 2-player Stackelberg differential game studied in this chapter is the first step towards defining and constructing approximate solutions to differential games with more complicated information

structures.

The task of monitoring a region using a team of agents equipped with sensors, *i.e.* optimal monitoring, is studied in Chapter 5. A general literature review on this topic is given in Section 5.2, before focusing on approaches that make use of the framework provided by game theory. In Section 5.4 the problem is formulated as a differential game. Although this problem formulation has several advantages, it is particularly notable that it allows for the use of *heterogeneous* agents. The problem is defined within the framework of differential games and a way of obtaining ad-hoc solutions for the optimal monitoring problem is then presented and the theory is illustrated by simulations.

In Chapter 6 mean-field games are introduced. Using the standard framework for mean-field games introduced by J.M. Lasry and P.L. Lions and by M.Y. Huang, P.E. Caines and R.P. Malhamé, a brief introduction and background to mean-field games is given in Section 6.2. In problems involving mean-field games it is common to include uncertainties and disturbances, for this reason the problems considered in Chapter 6 include elements of stochastic control theory. As a result the theory is somewhat different in nature to the preceding parts of the thesis. A background on stochastic control theory is not included in this thesis, and for more information on this the reader is, for example, referred to [34]. A specific class of scalar mean-field games and their solutions are then considered in Sections 6.3.1, 6.3.2 and 6.3.3. Here, we prove that for certain classes of mean-field games, obtaining solutions relies on solving a two-point boundary value problem. Simulations illustrating the theory are provided in Section 6.4.

Finally, the thesis is concluded in Chapter 7 in which a summary of the work and directions for future work are provided.

1.3 Notation

Standard notation has been adopted in this thesis, some of which is defined in this section and used throughout the remainder of this thesis. When new notation, not included in this section, is introduced these will be defined in the relevant parts of the thesis.

First, $\mathbb{R}$ denotes the set of real numbers, whereas $\mathbb{R}^+$ denotes the set of positive real numbers. Likewise, $\mathbb{C}$ denotes the complex plane and $\mathbb{C}^-$ denotes the left half of the complex plane. The set of positive natural numbers is denoted by $\mathbb{Z}^+$.

Let $x \in \mathbb{R}^n$, $n > 0$ and let $V(x)$ denote a function of the vector x . Suppose the function is continuously differentiable with respect to x . Then $\frac{\partial V}{\partial x}$ is used to denote the gradient of the function with respect to the vector x .

Consider a matrix $M \in \mathbb{R}^{n \times n}$ and let $M^\top$ denote its transpose. The spectrum of M is then denoted by $\sigma(M)$. Furthermore I and 0 are used to denote that identity and zero matrices, respectively. The matrix M is said to be positive definite if $x^\top M x > 0$, for all $x \in \mathbb{R}^n$. Similarly, M is said to be positive semi-definite if $x^\top M x \geq 0$, for all $x \in \mathbb{R}^n$. A (block) diagonal matrix is denoted by

$$\text{diag}(M_{11}, \dots, M_{NN}) = \begin{bmatrix} M_{11} & 0 & \dots & 0 \\ 0 & M_{22} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & 0 \end{bmatrix}.$$

Suppose $M = M^\top > 0$ and consider a vector $x \in \mathbb{R}^n$. The Euclidean norm of x is denoted by $\|x\|$, whereas $\|x\|_M$ denotes the Euclidean norm of the vector weighted by the matrix M , i.e. $\|x\|_M = (x^\top M x)^{\frac{1}{2}}$.

1.4 Published Results

The work presented in Chapter 3 and preliminary results have been published in [35–37] and the multi-agent collision avoidance problem discussed in Section 3.8 is considered in [38]. The results relating to Stackelberg differential games have been submitted for publication in [39]. The contributions related to *optimal monitoring* in Chapter 5 have been published in [40] and partly in [41]. The results relating to mean-field games in Chapter 6, and more specifically in Sections 6.3 and 6.4 have been published in [42, 43]. Some other results relating to mean-field games, which are not included in this thesis,

have been published in [44].

Chapter 2

Introduction and Preliminaries to Differential Game Theory

2.1 Introduction

Game theory, with its roots in von Neumann and Morgenstern's works during the late 1940's, is the study of multi-player decisions. Differential games, in turn, consider *multi-player decision making over time*. Since most situations call for making choices of some kind, differential games have a vast range of applications. Traditionally game theoretic methods play a significant role in the study of management and economics [45–47], and military and defense [31, 48]. They also appear in mathematical biology, for example in the study of the evolution of biological systems [49–51]. Furthermore, the agents in a multi-agent systems can be considered as players and thus, as will be seen in later sections, differential game theory is applicable to many problems involving multi-agent systems. Differential games also play an important role in the field of robust control, which deals with the study of dynamic systems and feedback control in the presence of uncertainties and disturbances [52–54]. For example, H_∞ optimal control, where controller synthesis with guaranteed stability properties in the presence of a disturbance signal is studied, boils down to a minimax optimisation problem and can be thought of as a *zero-sum* differential game where the controller seeks to minimise a cost which the

disturbance seeks to maximise [13]. For linear systems, the problem of mixed H_2/H_∞ control has been studied as a nonzero-sum game [14, 55]. Mixed H_2/H_∞ control for a class of nonlinear systems has been studied in [15]. In summary, there is no doubt that differential games play a central role in a diversity of applications.

As mentioned in Chapter 1, differential games can be considered as a generalisation of optimal control. In the latter, the problem lies in determining the best strategy for *one* player attempting to optimise a performance criterion subject to the dynamics of the state variable, whereas in the former, the problem is to determine the strategies for *several* players, each attempting to optimise its own performance criterion subject to common state dynamics. Thus differential games introduce the notion of *strategic behaviour* to control theory [6, 31, 32, 56–60]. A significant amount of work has been done in the field of zero-sum games, which consider the case in which the sum of the performance criteria of the players is zero, as, for example, is the case in a game of poker, a closed market with several buyers and sellers or an election [6, 10]. In particular two-player zero-sum games, in which the players have opposite performance criteria, *i.e.* a gain for one player implies a similar loss for its opponent, have been studied extensively by Isaacs [31]. The standard pursuit-evasion game with two participants is a classic example of such a differential game [60].

In zero-sum differential games a gain for one player implies a similar loss for its opponents and *vice versa* and, as a result, this class of games describes purely competitive situations [6, 31, 60]. However, in some situations, a gain for one player need not imply a loss for another. In certain scenarios involving decisions, one may encounter “win-win situations” which result in a positive outcome for all parties. Such situations, and many other scenarios, are not captured by the class of games that can be described as zero-sum. Non-zero sum differential games, on the other hand, regard the wide class of differential games, where there are several players, each attempting to optimise its own, *individual*, performance criterion, subject to the common state dynamics. Thus players are allowed to have different goals, which may or may not be competing [6, 32, 33, 58, 61]. As a result nonzero-sum differential games can be used to study and describe many different

scenarios that involve strategic decisions and consequently have a large variety of applications. For example, differential games that are not necessarily purely competitive appear in economics and management, defense, evolutionary biology, political science and social networks [31,45,47,49–51,62]. A well-known example of a nonzero-sum game is the Prisoner's Dilemma [6].

It is clear that differential games theory provides a powerful framework that can be useful to several engineering applications. Solving a differential game boils down to determining *equilibrium strategies* for all the players [6,31,32,58]. In the following two chapters we focus on *feedback Nash equilibrium strategies* for non-cooperative, nonzero-sum differential games. Determining these equilibrium strategies involves solving the Hamilton-Jacobi-Isaacs partial differential equations (PDEs) associated with the problem [6]. The Hamilton-Jacobi-Isaacs PDEs are a generalisation of the Hamilton-Jacobi-Bellman PDE encountered in optimal control. Closed-form solutions to PDEs are not, in general, readily found. Thus, solving PDEs is often the main obstacle when studying optimal control problems and, since solving a differential game requires solving coupled PDEs, the same challenge is encountered when seeking solutions for differential games. Therefore, although differential games theory provides a powerful framework that can be useful to several engineering applications, closed-form solutions may not be available making it necessary to seek approximate solutions of the Hamilton-Jacobi-Isaacs PDEs.

In [63] a method of constructing approximate solutions to optimal control problems using a dynamic feedback and the notion of a *algebraic $\bar{P}$ solution* has been developed. In Chapter 3 a somewhat similar approach is taken to obtain approximate solutions to a class of nonzero-sum differential games. However, in the case of N -player nonzero-sum differential games, the Hamilton-Jacobi-Isaacs equations consist of a system of N *coupled* PDEs, where N is the number of players participating in the game, in place of a single PDE encountered in optimal control problems and in H_∞ optimal control [6,32]. Thus, it is necessary to consider a notion somewhat different from the *algebraic $\bar{P}$ solution* of [63], namely, we introduce *algebraic $\bar{P}$ matrix solution*. Two methods of constructing dynamic control laws that satisfy N *coupled* partial differential inequalities are

derived. For optimal control problems, solving a partial differential inequality instead of the Hamilton-Jacobi-Bellman PDE corresponds to solving an optimal control problem with an additional running cost associated to the *negativity* gap in the inequality, so the *level of approximation* is directly quantifiable. In differential games, however, solving partial differential inequalities in place of the Hamilton-Jacobi-Isaacs equations is related to ϵ -Nash equilibria and the *level of approximation* is not directly quantifiable in the same way as is the case with optimal control problems.

The purpose of this chapter is to provide background information on differential games that serves as an introduction to the problems described in Chapter 3. With this in mind, differential games and their solutions in terms of *feedback Nash equilibrium strategies* are considered. First, a differential game and its solution are formally defined. A special class of *nonzero-sum* differential games, namely *linear-quadratic differential games*, is then described and finally the notion of ϵ -Nash equilibrium strategies is defined. To clarify the various concepts and notation 2-player differential games are considered before the attention is shifted to general N -player differential games. A more thorough background on differential game theory can be found in [6, 31, 32].

2.2 Problem Formulation

A differential game consists of N players, where $N \in \mathbb{Z}^+$. For the special case in which $N = 1$ the differential games becomes a problem of optimal control, as discussed in Section 2.1. We therefore focus our attention on differential games with $N \geq 2$ players. The problem of solving infinite-horizon nonzero-sum differential games in terms of feedback Nash equilibrium strategies is defined in this section.

2.2.1 2-Player Differential Games

Consider a system with state $x(t) \in \mathbb{R}^n$ and dynamics

$$\dot{x} = \tilde{f}(x, u_1, u_2), \quad (2.1)$$

where $\tilde{f}(x, u_1, u_2)$ is a mapping and $u_1(t) \in \mathbb{R}^{m_1}$ and $u_2(t) \in \mathbb{R}^{m_2}$, with $m_1 \leq n$ and $m_2 \leq n$, are the control actions, often referred to as the strategies, of players 1 and 2, respectively. Player 1 seeks to select its strategy $u_1(t)$ to minimise the cost functional

$$J_1(x(0), u_1, \dots, u_N, T) = \int_0^T q_1(x, u_1, u_2) dt + r_1(x(T)),$$

where $q_1(x, u_1, u_2)$ is a running cost and $r_1(x(T))$ is a terminal cost. Similarly, player 2 seeks to minimise the cost functional

$$J_2(x(0), u_1, \dots, u_N, T) = \int_0^T q_2(x, u_1, u_2) dt + r_2(x(T)),$$

where $q_2(x, u_1, u_2)$ and $r_2(x(T))$ are the running cost and terminal cost, respectively. This problem is known as a *2-player finite-horizon differential game*. The two players must determine their strategies u_1 and u_2 to minimise these cost functionals subject to the state dynamics (2.1).

In the limit as T approaches infinity the problem becomes a *2-player infinite-horizon differential game*. In some applications, for example in the setting of economics, the duration of a game may be long or even unknown. It is then common to consider these to be of infinite horizon [45]. In infinite-horizon differential games it is assumed that by means of their respective control strategies, players 1 and 2 seek to minimise cost functionals of the form

$$\begin{aligned} J_1(x(0), u_1, u_2) &= \int_0^\infty \tilde{q}_1(x, u_1, u_2) dt, \\ J_2(x(0), u_1, u_2) &= \int_0^\infty \tilde{q}_2(x, u_1, u_2) dt. \end{aligned} \tag{2.2}$$

Note that there is no terminal cost in the cost functionals for the infinite-horizon differential game and care must be taken to ensure that the cost functionals are bounded. Often this is ensured by imposing certain restrictions on the cost functionals, as in Chapter 3, or by introducing a so-called *discount factors* to the running costs, as in [45].

The cost functionals of the players need not be conflicting, *i.e.* the players may, but need not be, competing. For the special case in which the sum of the cost functionals of

the two players is zero, *i.e.* $q_1(x, u_1, u_2) = -q_2(x, u_1, u_2)$, a gain for player 1 implies an equal loss for player 2 and *vice versa*. Differential games that fall within this category are known as *2-player zero-sum differential games* [6, 31]. The general case in which the sum of the two cost functionals is not zero is known as a *2-player nonzero-sum differential game*. In what follows the notation $\mathcal{S} = \{u_1, u_2\}$ is used to denote the strategies u_1 and u_2 adopted by players 1 and 2, respectively.

In an optimal control problem the optimal control is such that the cost functional is minimised subject to the system dynamics. For a differential game, on the other hand, the concept of *optimality* is not as straight-forward and intuitive. Consider, for example, a two-player differential game and two sets of strategies $S_1 = \{u_1, u_2\}$ and $S_2 = \{w_1, w_2\}$ and suppose that the set of strategies S_1 is *favourable* for player 1 whereas the set S_2 is *favourable* for player 2. Clearly, in contrast to what is the case for optimal control problems, it is not straight-forward to determine which of the two sets of strategies is “better” than the other. Thus, different notions of solutions of differential games must be introduced.

The definition of optimality for differential games is not unique and several different solution concepts exist, the most common one being the so-called *Nash equilibrium strategies* [11]. We focus on *feedback Nash equilibrium strategies*, which are formalised in the following definitions. In what follows u_i is used to denote the feedback strategy of player i , *i.e.* $u_i(t) = u_i(x(t))$.

Definition 1. A pair of state feedback control strategies $\mathcal{S} = \{u_1, u_2\}$ is said to be *admissible* for the non-cooperative differential game (2.1), (2.2), if the zero-equilibrium of the system (2.1) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable.

Definition 2. The state feedback control strategies u_1^* and u_2^* are said to be *Nash equilibrium strategies* of player 1 and player 2, respectively, for the non-cooperative differential game (2.1), (2.2), if the pair of strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ is admissible and satisfies the inequalities

$$\begin{aligned} J_1(x(0), u_1^*, u_2^*) &\leq J_1(x(0), u_1, u_2^*) \\ J_2(x(0), u_1^*, u_2^*) &\leq J_2(x(0), u_1^*, u_2), \end{aligned} \tag{2.3}$$

for all admissible feedback strategy pairs $\mathcal{S}_1 = \{u_1, u_2^*\}$ and $\mathcal{S}_2 = \{u_1^*, u_2\}$ where $u_1 \neq u_1^*$ and $u_2 \neq u_2^*$.

The set of Nash equilibrium strategies $\mathcal{S}^*$ is referred to as the *feedback Nash equilibrium solution* of the differential game. As mentioned in Chapter 1 this solution concept assumes that all players have full access to information regarding the state x , the cost functionals of all the players and that the players announce their strategies simultaneously. This assumption is relaxed in Chapter 4 where a 2-player Stackelberg differential game is considered.

The Nash equilibrium solution of a differential game is such that if any one deviates from its Nash equilibrium strategy, while assuming all other players are rational, this results in a loss for the deviating player [32]. This does not, however, imply that the players cannot both perform better by adhering to different strategies. In fact, strategies that result in the best outcomes for both players, if they exist, are known as *Pareto optimal strategies*.

We now focus our attention on obtaining feedback Nash equilibrium solutions for nonzero-sum differential games with infinite horizon, *i.e.* we consider the following problem.

Problem 1. Consider the system (2.1) and the cost functionals (2.2). The problem of solving the 2-player, non-cooperative, differential game consists in determining a pair of admissible feedback strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ such that the inequalities (2.3) hold for all admissible pairs of feedback strategy $\mathcal{S}_1 = \{u_1, u_2^*\}$ and $\mathcal{S}_2 = \{u_1^*, u_2\}$ where $u_1 \neq u_1^*$ and $u_2 \neq u_2^*$.

As in optimal control one way of obtaining solutions to differential games is through dynamic programming [6, 32]. Players 1 and 2 are associated with value functions, $V_1(x(0), u_1, u_2)$ and $V_2(x(0), u_1, u_2)$, such that

$$V_1(x(0), u_1, u_2) = J_1(x(0), u_1, u_2),$$

$$V_2(x(0), u_1, u_2) = J_2(x(0), u_1, u_2).$$

Using the principles of dynamic programming it can be shown that the value functions satisfy

$$\min_{u_i} H_i \left(x, u_1, u_2, \frac{\partial V_i}{\partial x} \right) = 0, \quad (2.4)$$

where $i = 1, 2$ and $H_i(x, u_1, u_2, \lambda_i)$ is the Hamiltonian function of player i defined as

$$H_i(x, u_1, u_2, \lambda_i) = \tilde{q}_i(x, u_1, u_2) + \lambda_i^\top f(x(t), u_1(t), u_2(t)),$$

and λ_i is the costate of player i . Suppose $u_1^* = u_1 \left(x(t), \frac{\partial V_1}{\partial x}, \frac{\partial V_2}{\partial x} \right)$ and $u_2^* = u_2 \left(x(t), \frac{\partial V_1}{\partial x}, \frac{\partial V_2}{\partial x} \right)$ are the control strategies that achieve the minimum in (2.4) for $i = 1, 2$. It then follows that the value function $V_i(x_0, u_1, u_2)$ satisfies the PDE

$$H_i \left(x, u_1^*, u_2^*, \frac{\partial V_i}{\partial x} \right) = 0, \quad (2.5)$$

for $i = 1, 2$, subject to the system dynamics (2.1). Note that (2.5) hold for infinite-horizon problems only, as the partial derivatives of the value functions with respect to time, *i.e.* $\frac{\partial V_i}{\partial t}$, are nonzero and the Hamiltonians are functions of time for finite-horizon problems. Thus, obtaining solutions to the 2-player infinite-horizon nonzero-sum differential game, *i.e.* Problem 1 involves solving the coupled PDEs resulting from (2.5). Supposing stabilising solutions to these PDEs can be obtained, the Nash equilibrium strategies for players 1 and 2, which constitute the solution of Problem 1, are u_1^* and u_2^* , respectively.

2.2.2 N -player differential games

The concepts introduced for 2-player differential games can be extended to the general case in which there are N players in a straight-forward manner. Consider the N -player equivalent of (2.1), *i.e.* consider the dynamical system

$$\dot{x} = \tilde{f}(x(t), u_1(t), \dots, u_N(t)), \quad (2.6)$$

where $\tilde{f}(x, u_1, \dots, u_N)$ is a mapping and $u_i(t) \in \mathbb{R}^{m_i}$, with $m_i \leq n$, is the control strategy of player i , for $i = 1, \dots, N$. Furthermore, each of the players aims to select its control strategies to minimise its own cost functionals, which is of the form

$$J_i(x, u_1, \dots, u_N, T) = \int_0^T q_i(x, u_1, \dots, u_N) dt + r_i(x(T)),$$

for $i = 1, \dots, N$, where $q_i(x, u_1, \dots, u_N)$ is a running cost and $r_i(T)$ is a terminal cost, $i = 1, \dots, N$. This problem is known as the *N-player finite-horizon differential game*. Similarly to what was seen in Section 2.2.1, *N-player infinite-horizon differential games* considers the problem in which T approaches infinity and the cost functionals of the players are of the form

$$J_i = \int_0^\infty \tilde{q}_i(x, u_1, \dots, u_N) dt, \quad (2.7)$$

for $i = 1, \dots, N$

The special case in which the sum of the players cost functionals is zero is known as *N-player zero-sum differential games*, whereas the more general class of *N-player differential games* in which the sum of the costs is not zero is known as *N-player nonzero-sum differential games*. The latter is considered herein and, as with the 2-player differential games encountered in Section 2.2.1, we focus our attention on infinite-horizon differential games. The concepts of admissible strategies and feedback Nash equilibrium solutions for *N-player differential games* are natural extensions of Definitions 1 and 2 seen in the 2-player case.

Definition 3. The set of feedback control strategies $\mathcal{S} = \{u_1, \dots, u_N\}$ is said to be *admissible* for the non-cooperative differential game (2.6), (2.7) $i = 1, \dots, N$, if the zero-equilibrium of the system (2.6) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable.

Definition 4. The state feedback control strategies $u_1^*, \dots, u_N^*$ are said to be *Nash equilibrium strategies* of players $1 \dots N$, respectively, for the non-cooperative differential game (2.6), (2.7), if the set strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ is admissible and satisfies the inequalities

$$J_i(u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_i, \dots, u_N^*), \quad (2.8)$$

for all admissible feedback sets of strategies $S = \{u_1^*, \dots, u_i, \dots, u_N^*\}$, with $u_i \neq u_i^*$.

The set of feedback strategies S^* is said to be the *Nash equilibrium solution* of the N -player differential game. The infinite-horizon N -player nonzero-sum differential game is then defined as follows.

Problem 2. Consider the system (2.6) and the cost functionals (2.7). The problem of solving the N -player, non-cooperative differential game consists in determining a set of admissible feedback strategies $S^* = \{u_1^*, \dots, u_N^*\}$ such that the inequalities (2.8), $i = 1, \dots, N$, $j = 1, \dots, N$, $j \neq i$, hold for all admissible sets of strategies $S = \{u_1^*, \dots, u_i, \dots, u_N^*\}$ where $u_i \neq u_i^*$, for $i = 1, \dots, N$.

As in the two-player case solutions can be found by applying principles of dynamic programming. The Hamiltonian associated with player i is

$$H_i(x, u_1, \dots, u_N, \lambda_i) = \tilde{q}_N(x, u_1, \dots, u_N) + \lambda_i^\top f(x(t), u_1(t), \dots, u_N(t)),$$

where λ_i is the costate. Solving Problem 2 involves obtaining solutions to N coupled PDEs

$$H_i \left(u_1^*, \dots, u_N^*, \frac{\partial V_i}{\partial x} \right) = 0, \quad (2.9)$$

$i = 1, \dots, N$, subject to the state dynamics (2.6), where the Nash equilibrium strategy of player i , u_i^* , is the minimiser of the Hamiltonian of the player i . Consequently, solving N -player nonzero-sum differential games involves obtaining (stabilising) solutions for N coupled PDEs.

2.3 Linear-Quadratic Differential Games

In Chapter 3 a class of 2- and N -player nonlinear differential games is considered. To provide some insight prior to this, and to illustrate the background presented in Section 2.2, *linear-quadratic differential games* are discussed in this section. This class of differential games has been extensively studied in the literature, see for example [64]. In short, this is

the class of differential games in which the system satisfies linear dynamics and the cost functionals, which the players seek to minimise, are quadratic.

Linear-quadratic differential games are comparable to their counterpart in optimal control, namely linear-quadratic regulator problems. The optimal control for such a problem relies on the solution of an algebraic Riccati equation (ARE). In the case of linear-quadratic differential games, obtaining Nash equilibrium feedback strategies is not as straight-forward. The first complication is that Nash equilibrium solutions to linear quadratic differential games may be nonlinear, *i.e.* the feedback strategies u_i^* , $i = 1, \dots, N$, satisfying (2.8), if they exist, may be nonlinear functions of the state [64,65]. However, it is common to seek linear feedback strategies for linear-quadratic differential games.

As in [64,65], we restrict the solutions of linear-quadratic differential games to the class of linear feedback strategies. The rationale behind this choice is twofold. Firstly, since the dynamical system is linear it is natural to consider linear control laws that do not perturb the nature of the closed-loop system. Secondly, it is computationally appealing to determine such linear feedback strategies by solving systems of coupled matrix equations, or inequalities, in place of partial differential equations as in the genuinely nonlinear setting.

Among the class of linear feedback strategies, Nash equilibrium solutions to linear-quadratic differential games are obtained by solving a set of coupled AREs as discussed in this section. However, even when restricting ourselves to linear feedback strategies for linear-quadratic differential games, obtaining the Nash equilibrium solutions is not trivial, due to the nature of the coupled AREs. In what follows 2- and N -player linear-quadratic differential games are considered. Some remarks on the coupled AREs and their solutions are then given in Section 2.3.3.

2.3.1 2-player Linear-Quadratic Differential Games

Consider a 2-player differential game in which the system (2.6) is linear, namely the system dynamics are given by

$$\dot{x} = Ax + B_1 u_1 + B_2 u_2, \quad (2.10)$$

where x , u_1 , and u_2 are the state and the players' strategies as defined in Section 2.2, and $A \in \mathbb{R}^{n \times n}$, $B_1 \in \mathbb{R}^{n \times m_1}$ and $B_2 \in \mathbb{R}^{n \times m_2}$ are constant matrices. Furthermore, the players seek to minimise the following cost functionals, which are quadratic in the state variable and the control strategies,

$$\begin{aligned} J_1 &= \frac{1}{2} \int_0^\infty \left(x^\top Q_1 x + u_1^\top R_{11} u_1 + u_2^\top R_{12} u_2 \right) dt, \\ J_2 &= \frac{1}{2} \int_0^\infty \left(x^\top Q_2 x + u_1^\top R_{21} u_1 + u_2^\top R_{22} u_2 \right) dt, \end{aligned} \quad (2.11)$$

where $Q_1 \geq 0$, $Q_2 \geq 0$ and the weighting matrices satisfy $R_{11} = R_{11}^\top > 0$, $R_{22} = R_{22}^\top > 0$. Differential games of this form are known as *2-player linear-quadratic nonzero-sum differential games*.

Theorem 1. [32] Suppose $P_1 = P_1^\top$ and $P_2 = P_2^\top$ are such that $P_1 + P_2 > 0$ and satisfy the coupled AREs

$$\begin{aligned} &P_1 A + A^\top P_1 + Q_1 - P_1 B_1 R_{11}^{-1} B_1^\top P_1 - P_1 B_2 R_{22}^{-1} B_2^\top P_2 \\ &\quad - P_2 B_2 R_{22}^{-1} B_2^\top P_1 + P_2 B_2 R_{22}^{-1} R_{12} R_{22}^{-1} B_2^\top P_2 = 0, \\ &P_2 A + A^\top P_2 + Q_2 - P_2 B_2 R_{22}^{-1} B_2^\top P_2 - P_2 B_1 R_{11}^{-1} B_1^\top P_1 \\ &\quad - P_1 B_1 R_{11}^{-1} B_1^\top P_2 + P_1 B_1 R_{11}^{-1} R_{21} R_{11}^{-1} B_1^\top P_1 = 0. \end{aligned} \quad (2.12)$$

Then, the linear feedback strategies

$$u_1^* = -R_{11}^{-1} B_1^\top P_1 x, \quad u_2^* = -R_{22}^{-1} B_2^\top P_2 x, \quad (2.13)$$

are Nash equilibrium strategies for players 1 and 2, respectively, and $\mathcal{S}^* = \{u_1^*, u_2^*\}$ is a Nash equilibrium solution for the non-cooperative differential game in Problem 1 with dynamics (2.10) and cost functionals (2.11). $\diamond$

Proof: The proof then lies in showing that the strategies are consistent with (2.4) and (2.5) and that the set of strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ is admissible. The Hamiltonians of player 1 and 2 are

$$\begin{aligned} H_1(x, u_1, u_2, \lambda_1) &= \frac{1}{2}x^\top Q_1 x + \frac{1}{2}u_1^\top R_{11}u_1 + \frac{1}{2}u_2^\top R_{12}u_2 + \lambda_1^\top (Ax + B_1u_1 + B_2u_2), \\ H_2(x, u_1, u_2, \lambda_2) &= \frac{1}{2}x^\top Q_2 x + \frac{1}{2}u_1^\top R_{21}u_1 + \frac{1}{2}u_2^\top R_{22}u_2 + \lambda_2^\top (Ax + B_1u_1 + B_2u_2), \end{aligned}$$

respectively, and the control strategies u_1^* , minimising $H_1\left(x, u_1, u_2, \frac{\partial V_1}{\partial x}\right)$, and u_2^* , minimising $H_2\left(x, u_1, u_2, \frac{\partial V_2}{\partial x}\right)$, are

$$u_1^* = -R_{11}^{-1}B_1^\top \frac{\partial V_1}{\partial x}^\top, \quad u_2^* = -R_{22}^{-1}B_2^\top \frac{\partial V_2}{\partial x}^\top.$$

Furthermore, it follows from (2.12) that the quadratic value functions $V_1 = \frac{1}{2}x^\top P_1 x$ and $V_2 = \frac{1}{2}x^\top P_2 x$ satisfy the PDEs given by (2.5), i.e. the value functions satisfy

$$\begin{aligned} &\frac{\partial V_1}{\partial x} Ax + \frac{1}{2}x^\top Q_1 x - \frac{1}{2} \frac{\partial V_1}{\partial x} B_1 R_{11}^{-1} B_1^\top \frac{\partial V_1}{\partial x}^\top - \frac{\partial V_1}{\partial x} B_2 R_{22}^{-1} B_2^\top \frac{\partial V_2}{\partial x}^\top \\ &+ \frac{1}{2} \frac{\partial V_2}{\partial x} B_2 R_{22}^{-1} B_2^\top \frac{\partial V_2}{\partial x}^\top = 0, \\ &\frac{\partial V_2}{\partial x} Ax + \frac{1}{2}x^\top Q_2 x - \frac{1}{2} \frac{\partial V_2}{\partial x} B_2 R_{22}^{-1} B_2^\top \frac{\partial V_2}{\partial x}^\top - \frac{\partial V_2}{\partial x} B_1 R_{11}^{-1} B_1^\top \frac{\partial V_1}{\partial x}^\top \\ &+ \frac{1}{2} \frac{\partial V_1}{\partial x} B_1 R_{11}^{-1} B_1^\top \frac{\partial V_1}{\partial x}^\top = 0. \end{aligned}$$

By assumption $P_1 + P_2 > 0$, which implies that $W = V_1 + V_2 > 0$, for all $x \neq 0$, and the above PDEs imply that $\dot{W} < 0$, for all $x \neq 0$. Thus, admissibility of the set of strategies $\mathcal{S}^*$ follows from standard Lyapunov arguments. It follows that the linear feedback strategies (2.13) are Nash-equilibrium strategies for each of the two players and $\mathcal{S}^*$ is a Nash equilibrium solution for the differential game. $\square$

2.3.2 N-player Linear-Quadratic Differential Games

We now focus our attention on N -player linear-quadratic nonzero-sum differential games, which are a natural extension of the 2-player case.

Consider the dynamical system

$$\dot{x} = Ax + \sum_{i=1}^N B_i u_i, \quad (2.14)$$

and the cost functionals

$$J_i = \frac{1}{2} \int_0^\infty \left(\tilde{x}^\top Q_i x + u_i^\top R_{ii} u_i + \sum_{j=1, j \neq i}^N u_j^\top R_{ij} u_j \right) dt, \quad (2.15)$$

where $R_{ii} \geq 0$ and $i = 1, \dots, N$. As in the two-player case, the Nash equilibrium solution of this N -player differential game relies on the solution of the N coupled AREs as specified in the following theorem.

Theorem 2. [32] Suppose $P_i = P_i^\top$, $i = 1, \dots, N$ are such that $\sum_{i=1}^N P_i > 0$ and satisfy the coupled AREs

$$P_i A + A^\top P_i + Q_i - \sum_{j=1}^N \left(P_j B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j^\top P_j - P_i B_j R_{jj}^{-1} B_j^\top P_j - P_j B_j R_{jj}^{-1} B_j^\top P_i \right) = 0, \quad (2.16)$$

for $i = 1, \dots, N$. Then, the feedback strategy

$$u_i^* = R_{ii}^{-1} B_i^\top P_i x, \quad (2.17)$$

is a Nash equilibrium strategy for player i , for $i = 1, \dots, N$, and the set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ is a Nash equilibrium solution for the non-cooperative differential game in Problem 2 with dynamics (2.14) and cost functionals (2.15). $\diamond$

Proof: The proof follows the same steps as the proof of Theorem 1. Admissibility follows from the first part of the claim. Since equations (2.16) are satisfied, the quadratic value functions $V_i(x) = \frac{1}{2} x^\top P_i x$ satisfy the PDEs (2.9), $i = 1, \dots, N$ and it follows that (2.17) yields a Nash equilibrium strategy for player i and $\mathcal{S}^*$ is a Nash equilibrium solution for the differential game. $\square$

2.3.3 Some Remarks Concerning Linear-Quadratic Differential Games

As seen in Sections 2.3.1 and 2.3.2, when limiting the class of possible Nash equilibrium strategies for linear-quadratic differential games to the class of linear feedback strategies the equilibrium strategies rely on solutions of coupled AREs. Even in this case obtaining solutions is not straight-forward [66]. Linear-quadratic differential games and their associated coupled AREs have been studied extensively in the literature, see for example [66–70]. A comprehensive summary of results concerning solutions for both finite and infinite-horizon problems in terms of open- and closed-loop Nash equilibrium strategies is given in [64].

For finite-horizon problems it has been shown that open-loop Nash equilibrium strategies exist and are unique provided the time interval over which the game is played is sufficiently small [64, 67]. Similarly closed-loop Nash equilibrium solutions for finite-horizon differential games are studied in [71].

However, for infinite-horizon differential games the situation is somewhat more complicated. As seen in Sections 2.3.1 and 2.3.2, Nash equilibrium solutions can be found only if symmetric solutions to coupled AREs exist. Furthermore, these solutions have to be such that the zero equilibrium of the closed-loop system is stable. The existence of linear Nash strategies for infinite-horizon differential games is studied in [66], where it is shown that solutions exist provided the matrix describing the open-loop dynamics has a “sufficient degree of stability”. However, in general, for infinite-horizon problems, the coupled AREs may be such that many, one or no linear feedback Nash equilibrium solutions exist [64, 69]. For example, in [65, 72] a scalar, 2-player, infinite-horizon linear-quadratic game is considered and it is shown that even for this case solutions to the coupled AREs may be non-unique. Similarly, it is shown in [73] that for scalar, N -player, infinite-horizon linear-quadratic differential games the coupled AREs may have several solutions.

In [70] the computation of Nash equilibrium solutions for finite-horizon differential games is studied. Different iterative algorithms for computing stabilising solutions for infinite-horizon, linear-quadratic differential games exist, see for example [74]. How-

ever, the convergence properties of these rely heavily on the initial conditions and furthermore only provide one solution. In [73, 75] scalar linear quadratic differential games are considered and in Chapter 8 of [64] an algorithm for computing all feedback Nash equilibrium solutions for scalar linear-quadratic differential games is given.

From the above, it is clear that obtaining Nash equilibrium solutions, even for linear-quadratic differential games, may be difficult. Thus, it may in some cases be necessary to consider somewhat more relaxed concepts than that of Nash equilibrium solutions, which is introduced in the following section.

2.4 ϵ -Nash Equilibria

It is already apparent that obtaining solutions to differential games may be difficult, even for linear-quadratic differential games, and this will become even more apparent in Chapter 3. In general it may not be feasible to obtain closed-form solutions to the Hamilton-Jacobi-Isaacs PDEs, *i.e.* (2.5) in the 2-player case and (2.9), $i = 1, \dots, N$, in the N -player case, and even when it is possible, the computational complexity may be high. Thus, it is often of practical interest to consider a somewhat weaker solution concept, namely the notion of ϵ -Nash equilibria. The set of strategies which constitute a ϵ -Nash equilibrium can be interpreted as an approximation of a Nash equilibrium solution of a differential game and is a standard definition appearing in the literature, see for example Chapter 4 of [6].

For two-player differential games, the ϵ -equilibrium solution for a differential game is defined as follows.

Definition 5. The feedback strategies u_1^* and u_2^* are said to be ϵ -Nash equilibrium strategies for players 1 and 2, respectively, for Problem 1, if the pair of strategies $S^* = \{u_1^*, u_2^*\}$ is admissible and satisfies

$$\begin{aligned} J_1(u_1^*, u_2^*) &\leq J_1(u_1, u_2^*) + \epsilon, \\ J_2(u_1^*, u_2^*) &\leq J_2(u_1^*, u_2) + \epsilon, \end{aligned} \tag{2.18}$$

for some $\epsilon > 0$, for all admissible pairs of feedback strategies $\mathcal{S}_1 = \{u_1, u_2^*\}$ and $\mathcal{S}_2 = \{u_1^*, u_2\}$, with $u_1 \neq u_1^*$ and $u_2 \neq u_2^*$. The pair of strategies $\mathcal{S}^*$ is said to be an ϵ -Nash equilibrium of the differential game.

For N -player differential games ϵ -Nash equilibrium strategies are defined in a similar manner. Consider the N -player differential game in Problem 2.

Definition 6. The feedback strategies $u_1^*, \dots, u_N^*$ are said to be ϵ -Nash equilibrium strategies for players $1, \dots, N$, respectively, for Problem 2 if the set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ is admissible and satisfies

$$J_i(u_1^*, \dots, u_1^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_i, \dots, u_N^*) + \epsilon, \quad (2.19)$$

for some $\epsilon > 0$, and for any set of admissible strategies $\mathcal{S} = \{u_1^*, \dots, u_i, \dots, u_N^*\}$, with $u_i \neq u_i^*$, for $i = 1, \dots, N$. The set of strategies $\mathcal{S}^*$ is said to be an ϵ -Nash equilibrium of the differential game.

The sets of admissible strategies, $\mathcal{S}^*$ satisfying (2.18), in the two-player case, or (2.19), in the N -player case, are referred to as ϵ -Nash equilibrium solutions for Problems 1 and 2, respectively. Despite the fact that the concept of ϵ -Nash equilibrium clearly represents a relaxation, *i.e.* an approximation, of the stricter notion of Nash equilibrium, the former solution concept is of interest in practical cases since its computation may be significantly easier than that of a classical Nash equilibrium [6, 76].

The notion of ϵ -Nash equilibrium solution can be extended to different notions of solutions for differential games. In Chapter 4 it is extended to the case of Stackelberg equilibrium solutions in Chapter 4.

2.5 Conclusion

In this chapter a brief introduction and some background material on differential game theory have been given. It has been shown that obtaining feedback Nash equilibrium solutions generally relies on solving a system of coupled PDEs of the Hamilton-Jacobi-Isaacs type. For the class of linear-quadratic differential games, when considering linear

feedback strategies only, solutions rely on coupled AREs in place of the PDEs encountered in the more general setting. Furthermore, ϵ -Nash equilibrium solutions, which are a relaxation of Nash equilibrium solutions are defined. Since closed-form solutions to the Hamilton-Jacobi-Isaacs PDEs cannot in general be found, it is often necessary to solve differential games *approximately*. Using the background presented in this chapter two methods of constructing approximate solutions (in terms of feedback Nash equilibrium solutions) are developed in the next chapter.

Chapter 3

Constructive ϵ -Nash Equilibrium Solutions for Nonzero-Sum Differential Games

3.1 Introduction

General (infinite-horizon) nonzero-sum differential games and their solutions in terms of feedback Nash equilibrium strategies have been introduced in Chapter 2. As discussed therein, determining Nash equilibrium solutions for a differential game involves solving the Hamilton-Jacobi-Isaacs equations, *i.e.* equations (2.5), in the 2-player case, or equations (2.9) for $i = 1, \dots, N$, in the N -player case. As it is not, in general, possible to obtain closed-form solutions to PDEs it is often necessary to seek approximate solutions [76].

In this chapter a class of infinite-horizon, nonzero-sum differential games and their Nash equilibria are studied. Extending the results of [63], in which a method of approximately solving optimal control and robust control problems using a dynamic extension and the notion of *algebraic $\bar{P}$ solution* has been developed, a method of approximately solving the Hamilton-Jacobi-Isaacs PDEs is proposed. The Hamilton-Jacobi-Isaacs equation consists of a system of *coupled* PDEs, in place of a single PDE encountered in optimal control problems and in H_∞ control [6, 32]. Thus, the results in this chapter rely on a

notion somewhat different from the *algebraic $\bar{P}$ solution* of [63], namely, we introduce *algebraic $\bar{P}$ matrix solution*. Using these, two ways of constructing dynamic control laws satisfying partial differential inequalities are derived and it is demonstrated that these dynamic control laws approximate the solution of the original differential game. The problems considered are a subclass of the two-player differential games in Problem 1 and the N -player differential game in Problem 2, introduced in Chapter 2, namely, we focus on the case in which the system can be described by input-affine dynamics and the running costs associated with each agent is of a particular structure. The properties of the dynamics and running costs associated with the differential games considered herein are defined and for clarity the problems are stated again in this chapter.

The remainder of the chapter is structured as follows. In Section 3.2 we introduce the class of nonlinear differential games studied and recall some properties of the classical solutions, introduced in Chapter 2, for these games. Additionally the notions of α -admissible strategies and ϵ_α -Nash equilibrium strategies are introduced. Some properties of *linear-quadratic differential games* are discussed in Section 3.3, in which additionally solutions of specific matrix inequalities are related to the notion of ϵ_α -Nash equilibria. The aim of the remaining part of the chapter is then to extend the above results to the nonlinear setting avoiding the need for explicit solutions of partial differential equations or inequalities. Towards this end, in Section 3.4 we provide the notion of *algebraic $\bar{P}$ solution* that is instrumental for the construction of approximate solutions for the original class of nonlinear differential games. An approach exploiting this notion to obtain approximate solutions is presented in Section 3.5, for systems satisfying certain structural assumptions. In Section 3.6 another approach, which does not require these structural assumptions, is presented. Three numerical examples illustrating the theory are considered in Section 3.7. In the first example equilibrium strategies are known and it is demonstrated that the approximate methods proposed in this chapter result in outcomes closer to the Nash equilibrium outcomes than those resulting from solving the linear-quadratic approximation of the problem. The second example uses a Lotka-Volterra model. As a final example, the theory is then applied to a problem involving a multi-agent system in Section 3.8. In this section the problem of navigating a team of agents from given initial

positions to a desired end configuration while avoiding collisions is posed as a differential game. The problem is then solved using the results presented in 3.6. Finally, some conclusions are drawn in Section 7.2.

3.2 Nash and ϵ -Nash Equilibria for Non-cooperative differential games

Similarly to what has been done in Chapter 2, in this section we consider a dynamical system, with state x , and N players seeking to optimise different objectives, which may be conflicting, by selecting suitable control strategies, *i.e.* control actions [6]. The problem of solving a differential game (in terms of feedback Nash equilibria) is defined, as in Section 2.2, and properties of Nash equilibrium strategies and ϵ -Nash equilibrium strategies are recalled before the notion of ϵ_α -Nash equilibrium strategies is introduced. In particular, we consider differential games in which the system has input-affine dynamics and the cost functionals of each player satisfy certain properties as detailed in the following.

Consider the input-affine dynamics¹

$$\dot{x} = f(x) + \sum_{i=1}^N g_i(x)u_i, \quad (3.1)$$

and recall x , with $x(t) \in \mathbb{R}^n$, denotes the state of the system, which is affected by the actions of each player via the control input $u_i(t) \in \mathbb{R}^{m_i}$ associated to the i^{th} player, for $i = 1, \dots, N$.

Assumption 1. The origin of $\mathbb{R}^n$ is an equilibrium point of the vector field f , *i.e.* $f(0) = 0$.

Since $f(x)$ is smooth, Assumption 1 implies that there exists a continuous matrix-valued function $F : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ such that $f(x) = F(x)x$ for all $x \in \mathbb{R}^n$. Although $F(x)$ is possibly non-unique, $F(0)$ is unique. The above definition mimics the spirit of the so-called *apparent linearisation* [77], which is pursued in the State Dependent Riccati Equation (SDRE) approach to nonlinear optimal control problems [78]. Differently from the latter,

¹All mappings are assumed to be sufficiently smooth.

however, the methodology proposed herein yields a solution the *performances* of which – in terms for instance of asymptotic stability – are not affected by the choice of the matrix-valued function F .

Note that the system (3.1) may have several equilibria.

3.2.1 2-Player Problem Formulation

Consider first the case in which $N = 2$, *i.e.* the 2-player case. In this case the system (3.1) reduces to

$$\dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2. \quad (3.2)$$

The objectives of the two players are quantified by cost functionals, namely (2.2), which they seek to minimise. Consider the case in which these cost functionals are of the form

$$J_i(x(0), u_i, u_j) \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x) + \|u_i(t)\|^2 - \|u_j(t)\|^2 \right) dt, \quad (3.3)$$

for $i = 1, 2, j = 1, 2, j \neq i$. Note that J_1 and J_2 are parameterised with respect to the initial condition $x(0)$, and $q_i \in \mathcal{C}^r$, with $r \geq 2$, for $i = 1, 2$ are positive semidefinite running costs of the form $q_i(x) = x^\top Q_i(x)x$, $i = 1, 2$, without loss of generality, where the continuous matrix-valued functions $Q_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $i = 1, 2$, are positive semidefinite and symmetric for all $x \in \mathbb{R}^n$.

We consider running costs which satisfy the following assumption. The assumption is exploited later to ensure stability of the zero-equilibrium of the closed-loop system.

Assumption 2. The running costs $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are such that $q_1 + q_2$ is positive definite around the origin in $\mathbb{R}^n$.

Loosely speaking, each of the two players seeks to achieve a goal, quantified by its running cost, while minimising its own control action and maximising the effort of the other player. Before formally introducing the problem under investigation, it appears reasonable to define the set of control inputs that are considered *admissible*, in line with Definition 1, for the non-cooperative differential games dealt with in this section. As in

the previous chapter, the input u_i denotes the feedback strategy of player i , i.e. $u_i(t) = u_i(x(t))$.

Definition 7. A pair of state feedback control strategies $\mathcal{S} = \{u_1, u_2\}$ is said to be *admissible* for the non-cooperative differential game (3.2), (3.3), if the zero-equilibrium of the system (3.2) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable .

We define a stronger notion of admissibility in the following.

Definition 8. A pair of state feedback control strategies $\mathcal{S} = \{u_1, u_2\}$ is said to be α -*admissible*, with $\alpha > 0$, for the non-cooperative differential game (3.2), (3.3), if the zero-equilibrium of the system (3.2) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable and $\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of the closed-loop system around the origin.

The notion of α -admissible strategies is related to the notion of *discount factors*, commonly introduced in the context of infinite-horizon differential games: both enforce a minimum decay rate for the integrand function [45].

Problem 3. Consider the system (3.2) and the cost functionals (3.3). Solving the 2-player, non-cooperative differential game consists in determining a pair of admissible feedback strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ such that the inequalities $J_i^* \triangleq J_i(x(0), u_i^*, u_j^*) \leq J_i(x(0), u_i, u_j^*)$, $i = 1, 2, j = 1, 2, j \neq i$, hold for all admissible pairs of feedback strategies $\mathcal{S}_1 = \{u_1, u_2^*\}$ and $\mathcal{S}_2 = \{u_1^*, u_2\}$, where $u_1 \neq u_1^*$ and $u_2 \neq u_2^*$.

The set of strategies $\mathcal{S}^*$ constitutes a Nash equilibrium solution for the 2-player game, whereas (J_1^*, J_2^*) is the corresponding Nash equilibrium outcome, as remarked in Section 2.2.1 [6]. Suppose we are able to solve the system of two coupled Hamilton-Jacobi-Isaacs PDEs given by

$$\begin{aligned} \frac{\partial V_i}{\partial x} f(x) - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) - \frac{1}{2} \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \\ - \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top = 0, \end{aligned} \quad (3.4)$$

$i = 1, 2, j = 1, 2, j \neq i$, with $V_1(0) = 0, V_2(0) = 0$. The Nash equilibrium strategies for the two players are then given by

$$u_1^* = -g_1(x)^\top \frac{\partial V_1}{\partial x}^\top, \quad u_2^* = -g_2(x)^\top \frac{\partial V_2}{\partial x}^\top, \quad (3.5)$$

provided the system (3.2) in closed-loop with $\{u_1^*, u_2^*\}$ is locally asymptotically stable. Admissibility of the strategies in (3.5) is ensured if $W = V_1 + V_2 > 0$, for all $x \neq 0$.

Then

$$\dot{W} = \left(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial x} \right) \dot{x},$$

and (3.2) and (3.4) imply that $\dot{W} = -\frac{1}{2}(q_1(x) + q_2(x)) < 0$ along the trajectories of the closed-loop system. It follows that the pair of feedback strategies is admissible. Note that (3.4), $i = 1, 2$, are precisely the PDEs resulting from (2.5) for the class of differential games considered in this chapter.

For forward reference we introduce the linearised problem. The linearisation of the system (3.2) around the origin is

$$\dot{x} = Ax + B_1 u_1 + B_2 u_2, \quad (3.6)$$

with

$$A \triangleq \frac{\partial f}{\partial x} \Big|_{x=0} = F(0), \quad B_i \triangleq g_i(0), \quad (3.7)$$

where $i = 1, 2$. The running costs in (3.3) are replaced by the quadratic term $x^\top \bar{Q}_i x$, with $\bar{Q}_i \triangleq Q_i(0)$. The above describes a *linear-quadratic differential game* and determining Nash equilibrium solutions is not trivial even for such games, which may in general admit nonlinear feedback strategies. In the following, as in [64, 65] and as discussed in Section 2.3, we focus on the class of linear feedback strategies in the context of linear-quadratic differential games.

Among this class of strategies the Nash equilibrium strategies for the two players are given by

$$u_1^* = -B_1^\top \bar{P}_1 x, \quad u_2^* = -B_2^\top \bar{P}_2 x, \quad (3.8)$$

provided $\sigma(A - B_1 B_1^\top - B_2 B_2^\top \bar{P}_2) \subset \mathbb{C}^-$, where $\bar{P}_i, i = 1, 2$, are symmetric solutions of the *coupled AREs* [6, 32, 67]

$$\bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \bar{P}_i B_j B_j^\top \bar{P}_j - \left(\bar{P}_j B_j B_j^\top \bar{P}_i + \bar{P}_j B_j B_j^\top \bar{P}_j \right) + \bar{Q}_i = 0, \quad (3.9)$$

$i = 1, 2, j = 1, 2, j \neq i$. The control strategies (3.8) are admissible if $\bar{P}_1 + \bar{P}_2$ is positive definite. Then $W = x^\top (\bar{P}_1 + \bar{P}_2) x > 0$, for all $x \neq 0$, and (3.9) and Assumption 2 imply that $\dot{W} = -\frac{1}{2} x^\top (\bar{Q}_1 + \bar{Q}_2) x < 0$, for all $x \neq 0$. It follows that the pair of strategies (3.8) is admissible. The differential game defined by the linearised system given by (3.6) and (3.7), and the quadratic approximation of the costs is referred to as the *linear-quadratic approximation* of the problem.

The coupled AREs (3.9) are consistent with those introduced in Chapter 2, *i.e.* the equations are given by (2.12), with $R_{11} = R_{22} = I$ and $R_{12} = R_{21} = -I$. As discussed in Section 2.3.3, obtaining solutions for the coupled AREs (3.9) is generally not straightforward and consequently the (coupled) matrix equations (3.9) entail that a set of strategies which constitutes a Nash equilibrium solution may be difficult to determine even for linear time-invariant systems [66, 68–70, 79, 80]. Therefore, in the remainder of this section we recall the weaker notion of solution of non-cooperative differential games, namely the ϵ -Nash equilibrium strategies introduced in Definition 5. These partly motivate the results of the following sections.

Definition 9. Admissible feedback strategies u_1^* and u_2^* are said to be ϵ -Nash equilibrium strategies of players 1 and 2, respectively, of a 2-player, non-cooperative differential game if for the pair of strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ there exists a non-negative constant ϵ_{x_0} , parameterised with respect to the initial condition $x(0) = x_0$, such that $J_i(x_0, u_i^*, u_j^*) \leq J_i(x_0, u_i, u_j^*) + \epsilon_{x_0}$, for all admissible strategy pairs $\mathcal{S}_1 = \{u_1, u_2^*\}$ and $\mathcal{S}_2 = \{u_1^*, u_2\}$ with $u_1 \neq u_1^*$ and $u_2 \neq u_2^*$. The pair of strategies $\mathcal{S}_i^*$ is said to be an ϵ -Nash equilibrium of the differential game.

Finally, another notion of solution is defined as follows.

Definition 10. Admissible feedback strategies u_1^* and u_2^* are said to be ϵ_α -Nash equilibrium strategies for players 1 and 2, respectively, of a 2-player, non-cooperative differen-

tial game if for the pair of strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$ there exists a non-negative constant $\epsilon_{x_0, \alpha}$, parameterised with respect to the initial condition $x(0) = x_0$ and $\alpha > 0$, such that $J_i(x_0, u_i^*, u_j^*) \leq J_i(x_0, u_i, u_j^*) + \epsilon_{x_0, \alpha}$, for all α -admissible strategy pairs $\mathcal{S}_i = \{u_i, u_j^*\}$, with $u_i \neq u_i^*$ $i = 1, 2, j = 1, 2, j \neq i$. The pair of strategies $\mathcal{S}_i^*$ is said to be an ϵ_α -Nash equilibrium of the differential game.

Remark 1. Alternatively to the cost functionals (3.3) different weights on the control inputs u_i , $i = 1, 2$, can be considered, similarly to what has been done in [66]. In this scenario, sufficient conditions for admissibility of the Nash equilibrium solution, $\mathcal{S}^* = \{u_1^*, u_2^*\}$ are stated in the following result. $\blacktriangle$

Proposition 1. Consider system (3.2) and the 2-player differential game where the two players seek to minimise the cost functionals

$$J_i \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x(t)) + u_i(t)^\top R_{ii} u_i(t) - u_j(t)^\top R_{ij} u_j(t) \right) dt,$$

with $R_{ii} > 0$, $i = 1, 2, j = 1, 2, j \neq i$. The Hamilton-Jacobi-Isaacs equations associated with the differential game are

$$\begin{aligned} & \frac{\partial V_i}{\partial x} f(x) - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) R_{ii}^{-1} g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) - \frac{\partial V_i}{\partial x} g_j(x) R_{jj}^{-1} g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \\ & - \frac{1}{2} \frac{\partial V_j}{\partial x} g_j(x) R_{jj}^{-1} R_{ij} R_{jj}^{-1} g_j(x)^\top \frac{\partial V_j}{\partial x}^\top = 0 \end{aligned} \quad (3.10)$$

$i = 1, 2, j = 1, 2, j \neq i$, with $V_1(0) = 0$ and $V_2(0) = 0$. Suppose Assumption 2 holds. Then the strategies (3.5) are admissible provided $V_1 + V_2 > 0$, for all $x \neq 0$, $R_{11} \geq R_{21}$ and $R_{22} \geq R_{12}$. $\diamond$

Proof: Taking $W = V_1 + V_2$ as a candidate Lyapunov function, it follows from (3.10) that for all $x \neq 0$, $\dot{W} \leq -\frac{1}{2}(q_1 + q_2) < 0$ holds provided $R_{11} \geq R_{21}$ and $R_{22} \geq R_{12}$. $\square$

3.2.2 N-Player Problem Formulation

Consider now the general case in which the game involves $N > 2$ players. The notions and definitions introduced for this scenario are direct extensions of those introduced for

the case in which $N = 2$, but are presented for completeness. Similarly to the 2-player case, the objectives of each of the players are quantified by cost functionals, namely (2.7), $i = 1, \dots, N$, which the players seek to minimise. These cost functionals are of the form

$$J_i(x(0), u_1, \dots, u_N) \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x(t)) + \|u_i(t)\|^2 - \sum_{j=1, j \neq i}^N \|u_j(t)\|^2 \right) dt, \quad (3.11)$$

parameterised with respect to the initial condition $x(0)$ and the positive semidefinite running costs $q_i \in \mathcal{C}^r$, $i = 1, \dots, N$, of the form $q_i(x) = x^\top Q_i(x)x \geq 0$, where the continuous matrix-valued functions $Q_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $i = 1, \dots, N$, are positive semidefinite and symmetric for all $x \in \mathbb{R}^n$.

Assumption 3. The running costs $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are such that the function $\sum_{i=1}^N q_i(x)$ is positive definite around the origin in $\mathbb{R}^n$.

As in the 2-player case, Assumption 3 is exploited to ensure stability of the zero-equilibrium of the closed-loop system.

Similarly to Definitions 7 and 8, in this setting the set of *admissible* and α -*admissible* control strategies are defined. Recall first Definition 3, namely α -admissible strategies, for differential games described by the dynamics (3.1) and the cost functionals (3.11), $i = 1, \dots, N$.

Definition 11. The set of state feedback control strategies $\mathcal{S} = \{u_1, \dots, u_N\}$ is said to be *admissible* for the non-cooperative differential game if the zero-equilibrium of the system (3.1) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable.

The stronger notion of α -*admissible* strategies is a direct extension of Definition 8.

Definition 12. The set of state feedback control strategies $\mathcal{S} = \{u_1, \dots, u_N\}$ is said to be α -*admissible* for the non-cooperative differential game if the zero-equilibrium of the system (3.1) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable and $\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of the closed-loop system around the origin.

Problem 4. Consider the system (3.1) and the cost functionals (3.11), $i = 1, \dots, N$. The problem of solving the N -player non-cooperative differential game consists in determining a set of admissible feedback strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ such that $J_i^* \triangleq$

$J_i(x(0), u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(x(0), u_1^*, \dots, u_i, \dots, u_N^*)$, for all admissible sets of feedback strategies $\mathcal{S}_i = \{u_1^*, \dots, u_i, \dots, u_N^*\}$, where $u_i \neq u_i^*$ and $i = 1, \dots, N$.

As in the 2-player case, it follows from Definition 4 that the set of strategies $\mathcal{S}^*$ constitutes a Nash equilibrium solution for the N -player game, whereas $(J_1^*, \dots, J_N^*)$ is the corresponding Nash equilibrium outcome. Suppose we are able to solve the system of coupled Hamilton-Jacobi-Isaacs PDEs given by

$$\begin{aligned} \frac{\partial V_i}{\partial x} f(x) - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) \\ - \frac{1}{2} \sum_{j=1, j \neq i}^N \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top - \sum_{j=1, j \neq i}^N \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top = 0, \end{aligned} \quad (3.12)$$

with $V_i(0) = 0$. Then the Nash equilibrium strategy for each player is provided by the static state feedback

$$u_i^* = -g_i(x)^\top \frac{\partial V_i}{\partial x}^\top, \quad (3.13)$$

for $i = 1, \dots, N$, provided the set of strategies $\mathcal{S}^* = \{u_i^*, \dots, u_N^*\}$ is admissible. Similarly to the 2-player case, suppose that $W = \sum_{i=1}^N V_i > 0$, for all $x \neq 0$. It follows from (3.1) and (3.12), $i = 1, \dots, N$, that $\dot{W} = -\frac{1}{2} \sum_{i=1}^N q_i$ and, consequently, the set of strategies $\mathcal{S}^*$ is admissible.

The linearised problem is defined by the linear approximation of the system (3.1) about the origin, *i.e.*

$$\dot{x} = Ax + \sum_{i=1}^N B_i u_i, \quad (3.14)$$

with A and $B_1, \dots, B_N$ given by (3.7), with $i = 1, \dots, N$, and the quadratic approximations of the running costs, where the running costs in (3.11), $i = 1, \dots, N$, are replaced by the quadratic term $x^\top \bar{Q}_i x$, with $\bar{Q}_i \triangleq Q_i(0)$. Note that in the linearised case the set of admissible strategies, for both the 2-player case and the N -player case, coincide with the set of linear state feedbacks $u_i = K_i x$, $i = 1, \dots, N$, which are such that all the eigenvalues of the matrix $A + \sum_{i=1}^N B_i K_i$ have negative real parts. Within the class of linear feedback strategies the Nash equilibrium strategy for each player is provided by the static state feedback $u_i^* = -B_i^\top \bar{P}_i x$, for $i = 1, \dots, N$, where $\bar{P}_i$, $i = 1, \dots, N$, provided the

set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ is admissible, where $\bar{P}_i, i = 1, \dots, N$, are symmetric solutions of the N coupled AREs

$$\begin{aligned} \bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \sum_{j=1, j \neq i}^N \bar{P}_i B_j B_j^\top \bar{P}_j - \sum_{j=1, j \neq i}^N \bar{P}_j B_j B_j^\top \bar{P}_i \\ - \sum_{j=1, j \neq i}^N \bar{P}_j B_j B_j^\top \bar{P}_j + \bar{Q}_i = 0, \end{aligned} \quad (3.15)$$

for $i = 1, \dots, N, j = 1, \dots, N$ and $j \neq i$. As in the 2-player case, admissibility of the set of feedback strategies $\mathcal{S}^*$ follows if $W = x^\top \left(\sum_{i=1}^N \bar{P}_i \right) x > 0$ and $\dot{W} = \frac{1}{2} x^\top \left(\sum_{i=1}^N \bar{Q}_i \right) x < 0$, for all $x \neq 0$.

As discussed in the 2-player case, (3.15) is consistent with (2.17) with $R_{ii} = I$ and $R_{ij} = -I$, for $i = 1, \dots, N, j = 1, \dots, N, j \neq i$, and solving coupled AREs of this form is not generally straightforward. Thus, it is of interest to recall the definition of ϵ -Nash equilibria once again, and extend the notion of ϵ_α -Nash equilibria in Definition 10 to the more general case in which $N > 2$.

Definition 13. Admissible feedback strategies $u_1^*, \dots, u_N^*$ are said to be ϵ -Nash equilibrium strategies of players $1, \dots, N$, respectively, of a non-cooperative differential game if for the set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ there exists a non-negative constant ϵ_{x_0} , parameterised with respect to the initial condition $x(0) = x_0$, such that $J_i(x_0, u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(x_0, u_1^*, \dots, u_i, \dots, u_N^*) + \epsilon_{x_0}$, for all sets of admissible strategies $\mathcal{S}_i = \{u_1^*, \dots, u_i, \dots, u_N^*\}$, where $u_i \neq u_i^*$ and $i = 1, \dots, N$. The set of strategies $\mathcal{S}^*$ is said to be an ϵ -Nash equilibrium of the differential game.

Definition 14. Admissible feedback strategies $u_1^*, \dots, u_N^*$ are said to be ϵ_α -Nash equilibrium strategies of players $1, \dots, N$, respectively, of a non-cooperative differential game if for the set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ there exists a non-negative constant $\epsilon_{x_0, \alpha}$, parameterised with respect to the initial condition $x(0) = x_0$ and $\alpha > 0$, such that $J_i(x_0, u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(x_0, u_1^*, \dots, u_i, \dots, u_N^*) + \epsilon_{x_0, \alpha}$, for all sets of α -admissible strategies $\mathcal{S}_i = \{u_1^*, \dots, u_i, \dots, u_N^*\}$, where $u_i \neq u_i^*$ and $i = 1, \dots, N$. The set of strategies $\mathcal{S}^*$ is said to be an ϵ_α -Nash equilibrium of the differential game.

3.3 Linear-Quadratic Non-Cooperative Differential Games

The objective of this section is to present results that relate solutions of specific (matrix) inequalities to the concept of ϵ_α -Nash equilibrium in the simpler, and explanatory, case of linear-quadratic differential games². These results are then extended to the nonlinear setting by exploiting a tool introduced in Section 3.4 which allows to replace partial differential inequalities with matrix algebraic inequalities.

Proposition 2. Consider the linear system (3.6) and the cost functionals (3.3), with $q_i(x) = x^\top \bar{Q}_i x$, $i = 1, 2$. Consider the class of linear feedback strategies and suppose that $\bar{P}_i = \bar{P}_i^\top$, $i = 1, 2$, are such that $\bar{P}_1 + \bar{P}_2 > 0$ and satisfy the coupled Riccati *inequalities*

$$\bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \bar{P}_i B_j B_j^\top \bar{P}_j - \left(\bar{P}_j B_j B_j^\top \bar{P}_i + \bar{P}_j B_j B_j^\top \bar{P}_j \right) + \bar{Q}_i \leq 0, \quad (3.16)$$

$i = 1, 2, j = 1, 2, j \neq i$. Then the pair of strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$, with $u_1^* = -B_1^\top \bar{P}_1 x$ and $u_2^* = -B_2^\top \bar{P}_2 x$, is admissible and yields an ϵ_α -Nash equilibrium, for any $\alpha > 0$, of the non-cooperative differential game. $\diamond$

Proof: If the matrices $\bar{P}_1, \bar{P}_2$ satisfy the inequalities (3.16), then there exist positive semidefinite matrices $\Upsilon_i = \Upsilon_i^\top \geq 0$, $i = 1, 2$, such that

$$\bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \bar{P}_i B_j B_j^\top \bar{P}_j - \left(\bar{P}_j B_j B_j^\top \bar{P}_i + \bar{P}_j B_j B_j^\top \bar{P}_j \right) + \bar{Q}_i + \Upsilon_i = 0, \quad (3.17)$$

$i = 1, 2, j = 1, 2, j \neq i$. By the equations (3.17) and recalling the definition of the coupled AREs (3.9), it follows that $\mathcal{S}^*$ represents a Nash equilibrium for the system (3.6) with respect to the modified cost functionals

$$\tilde{J}_i(x_0, u_i, u_j) \triangleq \frac{1}{2} \int_0^\infty \left(x(t)^\top (\bar{Q}_i + \Upsilon_i) x(t) + \|u_i(t)\|^2 - \|u_j(t)\|^2 \right) dt, \quad (3.18)$$

$i = 1, 2, j = 1, 2, j \neq i$, with $x(0) = x_0$. Note that $W = x^\top (\bar{P}_1 + \bar{P}_2) x > 0$, for all $x \neq 0$, and $\dot{W} = -\frac{1}{2} x^\top (\bar{Q}_1 + \Upsilon_1 + \bar{Q}_2 + \Upsilon_2) x < 0$, for all $x \neq 0$, hence the pair of strategies

²The results are stated and proved in the 2-player case: the extension to the case of N players is straightforward.

S^* is admissible. By definition of Nash equilibrium $\tilde{J}_i(x_0, u_i^*, u_j^*) \leq \tilde{J}_i(x_0, \hat{u}_i, u_j^*)$ for any admissible pair $\hat{S}_i = \{\hat{u}_i, u_j^*\}$, where $\hat{u}_i \neq u_i^*$ and u_j^* , with $i = 1, 2, j = 1, 2, j \neq i$. Then, note that $\tilde{J}_i(x_0, u_i^*, u_j^*) = \frac{1}{2} \int_0^\infty \left(x^*(t)^\top (\bar{Q}_i + \Upsilon_i) x^*(t) + \|u_i^*(t)\|^2 - \|u_j^*(t)\|^2 \right) dt$ and $\tilde{J}_i(x_0, \hat{u}_i, u_j^*) = \frac{1}{2} \int_0^\infty \left(\hat{x}(t)^\top (\bar{Q}_i + \Upsilon_i) \hat{x}(t) + \|\hat{u}_i(t)\|^2 - \|u_j^*(t)\|^2 \right) dt$, where $x^*(t)$ is the solution of system (3.6) in closed-loop with S^* and $x^*(0) = x_0$, whereas $\hat{x}(t)$ denotes the solution of the system $\dot{\hat{x}} = (A - B_j B_j^\top \bar{P}_j) \hat{x} + B_i \hat{u}_i$, $\hat{x}(0) = x_0$, for any arbitrary α -admissible pair $\hat{S}_i = \{\hat{u}_i, u_j^*\}$. Since the matrices Υ_i are positive semidefinite for $i = 1, 2$, it follows that $J_i(x_0, \hat{u}_i, u_j^*) \leq \tilde{J}_i(x_0, \hat{u}_i, u_j^*)$ and furthermore $\tilde{J}_i(x_0, \hat{u}_i, u_j^*) = J_i(x_0, \hat{u}_i, u_j^*) + \frac{1}{2} \int_0^\infty \hat{x}(t)^\top \Upsilon_i \hat{x}(t) dt$, hence

$$J_i(x_0, u_i^*, u_j^*) \leq J_i(x_0, \hat{u}_i, u_j^*) + \frac{1}{2} \int_0^\infty \hat{x}(t)^\top \Upsilon_i \hat{x}(t) dt, \quad (3.19)$$

for any α -admissible pair $\hat{S}_i$. Note that, by definition of α -admissible pairs, the second term in the right-hand side of (3.19) can be upper-bounded, for any initial condition, by a constant. Let K_i be such that $\hat{S}_i$ with $\hat{u}_i = K_i x$ is α -admissible and yields the maximum value of the integral³. Then, the right-hand side of (3.19) can be written in the form $J_i(x_0, \hat{u}_i, u_j^*) + \varepsilon_{x_0, \alpha}$, where $\varepsilon_{x_0, \alpha} \geq 0$ is parametrised in x_0 and α , and it is independent of the strategy $\hat{u}_i$. The second term of the right-hand side of inequality (3.19) can be written as

$$\frac{1}{2} \int_0^\infty \hat{x}(t)^\top \Upsilon_i \hat{x}(t) dt = \frac{1}{2} x_0^\top P_{i, \epsilon} x_0, \quad (3.20)$$

where $P_{i, \epsilon} = P_{i, \epsilon}^\top > 0$ solves $P_{i, \epsilon} A_{K_i} + A_{K_i}^\top P_{i, \epsilon} + \Upsilon_i = 0$. In fact, consider the function $V(x) = \frac{1}{2} x^\top P_{i, \epsilon} x$ the time derivative of which is given by $\dot{V}(x) = \frac{1}{2} x^\top P_{i, \epsilon} A_{K_i} x + \frac{1}{2} x^\top A_{K_i}^\top P_{i, \epsilon} x = -\frac{1}{2} x^\top \Upsilon_i x$, where $A_{K_i} = (A - B_j B_j^\top \bar{P}_j - B_i K_i)$. Integrating both sides of the above equation from zero to infinity yields $-\frac{1}{2} \int_0^\infty x(t)^\top \Upsilon_i x(t) dt = \frac{1}{2} \int_0^\infty \dot{V}(x(t)) dt = V(x(\infty)) - V(x_0)$, which proves the claim, since $\hat{S}_i$ is α -admissible, hence $V(x(\infty)) = 0$. The proof is concluded noting that, by (3.19), $J_i(x_0, u_i^*, u_j^*) \leq J_i(x_0, \hat{u}_i, u_j^*) + \frac{1}{2} x_0^\top P_{i, \epsilon} x_0$. This implies $J_i(x_0, u_i^*, u_j^*) \leq J_i(x_0, \hat{u}_i, u_j^*) + \epsilon_{x_0}(\alpha)$, where $\epsilon_{x_0}(\alpha) = \max_i \{x_0^\top P_{i, \epsilon} x_0\}$, $i = 1, 2$. $\square$

Proposition 2 implies that solving algebraic Riccati inequalities, which represents

³Such gain matrix K_i always exists by α -admissibility.

a substantial simplification with respect to the corresponding equations, yields a reasonable approximation, as detailed in Definition 8, of the concept of Nash equilibrium. In [81], for example, solvability criteria for algebraic Riccati equalities and inequalities, which appear in H_∞ optimal control, are discussed and it is shown that under some conditions algebraic Riccati inequalities may be solvable even though the corresponding AREs are not.

It might appear possible, in some circumstances, to render the second term of the right-hand side of (3.19) arbitrarily small by selecting K_i such that the eigenvalues of A_{K_i} have arbitrarily large, negative, real parts. However, the choice of the matrix K_i , hence of the control input $\hat{u}_i$, affects also the first term of the right-hand side of (3.19). Indeed, there exists a minimum of the above *trade-off*, which is characterized in the following proposition, in which Υ_i is given by (3.17).

Proposition 3. Among all admissible pairs $\hat{S}_i = \{\hat{u}_i, u_j^*\}$ the minimum of the right-hand side of inequality (3.19) is achieved for $\hat{u}_i = -B_i^\top \bar{P}_\epsilon x$, where $\bar{P}_\epsilon = \bar{P}_\epsilon^\top$ is the positive definite solution of $(A - B_j B_j^\top \bar{P}_j)^\top \bar{P}_\epsilon + \bar{P}_\epsilon (A - B_j B_j^\top \bar{P}_j) - \bar{P}_\epsilon B_i B_i^\top \bar{P}_\epsilon + \bar{Q}_i + \Upsilon_i = 0$, and the minimum is equal to $\frac{1}{2} x_0^\top \bar{P}_\epsilon x_0 - \frac{1}{2} \int_0^\infty \|u_j^*(t)\|^2 dt = \frac{1}{2} x_0^\top (P_\epsilon + P_{j,K_i}) x_0$, where P_{j,K_i} satisfies $P_{j,K_i} A_{K_i} + A_{K_i}^\top P_{j,K_i} + \bar{P}_j B_j B_j^\top \bar{P}_j = 0$. $\diamond$

Proof: The claim can be proved by applying standard arguments of optimal linear-quadratic control, noting that the problem can be recast into a classical optimal control problem with respect to the cost functional

$$\frac{1}{2} \int_0^\infty \left(\hat{x}(t)^\top (\bar{Q}_i + \Upsilon_i) \hat{x}(t) + \|\hat{u}_i(t)\|^2 \right) dt, \quad (3.21)$$

subject to the dynamical constraint $\dot{\hat{x}} = (A - B_j B_j^\top \bar{P}_j) \hat{x} + B_i \hat{u}_i$ and the initial condition $\hat{x}(0) = x_0$. $\square$

Remark 2. When the margin of stability, quantified in terms of the parameter α , is large then $\epsilon_{x_0, \alpha}$ is small. Conversely, when the margin of stability approaches zero, $\epsilon_{x_0, \alpha}$ may become unbounded. $\blacktriangle$

3.4 Algebraic $\bar{P}$ Matrix Solution

As mentioned in Chapter 2, finding closed-form solutions to the coupled PDE's (3.12), with $i = 1, \dots, N$ (or (3.4) in the 2-player case) is not generally possible and it is therefore often necessary to settle for approximate solutions. The approaches provided in this chapter rely on the notion of *algebraic $\bar{P}$ matrix solution*. As mentioned in the introduction, a similar notion has been introduced in [63, 82] for optimal/robust control problems. However, since differential games are inherently different from optimal control problems, in that they involve *coupled* PDEs, a somewhat different notion is employed in this case. This is explained in detail for the 2-player case and for the more general N -player case in the following.

3.4.1 2-Player Algebraic $\bar{P}$ Matrix Solution

We introduce first the notion of *algebraic $\bar{P}$ matrix solution* for the 2-player case, *i.e.* the non-cooperative differential game with two players attempting to minimise individual cost functionals subject to common state dynamics.

Definition 15. Consider the system (3.2) and the cost functionals (3.3). Let $\Sigma_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $\Sigma_i(0) \geq 0$, $i = 1, 2$, be matrix-valued functions such that $\Sigma_i(x) = \Sigma_i(x)^\top > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and let $\bar{\Sigma}_i = \Sigma_i(0)$. The $\mathcal{C}^1$ matrix-valued functions $P_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $P_i(x) = P_i(x)^\top$, $i = 1, 2$, are said to be an $\mathcal{X}$ -*algebraic $\bar{P}$ matrix solution*⁴ of the equations (3.4) provided the following conditions hold⁵.

- (i) For all $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $i = 1, 2$, $j \neq i$,

$$\begin{aligned} & P_i(x)F(x) + F(x)^\top P_i(x) - P_i(x)g_i(x)g_i(x)^\top P_i(x) + Q_i(x) \\ & - (P_j(x) + P_i(x))g_j(x)g_j(x)^\top P_j(x) - P_j(x)g_j(x)g_j(x)^\top P_i(x) + \Sigma_i(x) = 0, \end{aligned} \quad (3.22)$$

- (ii) $P_i(0) = \bar{P}_i$, such that $\bar{P}_1 + \bar{P}_2 > 0$, with $\bar{P}_i$, $i = 1, 2$, the symmetric solutions of the

⁴Provided the set $\mathcal{X}$ contains the origin.

⁵Recall that $F(x)$ is a matrix-valued function which is such that $f(x) = F(x)x$.

coupled Riccati equations

$$\bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \bar{P}_i B_j B_j^\top \bar{P}_j - \left(\bar{P}_j B_j B_j^\top \bar{P}_i + \bar{P}_j B_j B_j^\top \bar{P}_j \right) + \bar{Q}_i + \bar{\Sigma}_i = 0. \quad (3.23)$$

If $\mathcal{X} = \mathbb{R}^n$, then P_1 and P_2 are said to be an *algebraic $\bar{P}$ matrix solution*.

3.4.2 N -Player Algebraic $\bar{P}$ Matrix Solution

Consider now the general case in which the differential game consists of N participating players. The notion of *algebraic $\bar{P}$ matrix solution* for this setting is a natural extension of the 2-player case as defined below.

Definition 16. Consider the system (3.1) and the cost functionals (3.11) with $i = 1, \dots, N$. Let $\Sigma_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $\Sigma_i(0) \geq 0$, $i = 1, \dots, N$, be matrix-valued functions such that $\Sigma_i(x) = \Sigma_i(x)^\top > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and let $\bar{\Sigma}_i = \Sigma_i(0)$. The $\mathcal{C}^1$ matrix-valued function $P_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $P_i(x) = P_i(x)^\top$, $i = 1, \dots, N$, are said to be an *$\mathcal{X}$ -algebraic $\bar{P}$ matrix solution*⁵ of the equations (3.4) provided the following conditions hold.

- (i) For all $x \in \mathcal{X} \subseteq \mathbb{R}^n$ and $i = 1, \dots, N$,

$$\begin{aligned} & P_i(x)F(x) + F(x)^\top P_i(x) - P_i(x)g_i(x)g_i(x)^\top P_i(x) + Q_i(x) \\ & - (P_j(x) + P_i(x))g_j(x)g_j(x)^\top P_j(x) - P_j(x)g_j(x)g_j(x)^\top P_i(x) + \Sigma_i(x) = 0, \end{aligned} \quad (3.24)$$

- (ii) $P_i(0) = \bar{P}_i$, such that $\sum_{i=1}^N \bar{P}_i > 0$, with $\bar{P}_i$ the symmetric solution of the *coupled Riccati equations*

$$\begin{aligned} & \bar{P}_i A + A^\top \bar{P}_i - \bar{P}_i B_i B_i^\top \bar{P}_i - \sum_{j=1, j \neq i}^N \bar{P}_i B_j B_j^\top \bar{P}_j - \sum_{j=1, j \neq i}^N \bar{P}_j B_j B_j^\top \bar{P}_i \\ & - \sum_{j=1, j \neq i}^N \bar{P}_j B_j B_j^\top \bar{P}_j + \bar{Q}_i + \bar{\Sigma}_i = 0, \end{aligned} \quad (3.25)$$

for $i = 1, \dots, N$.

If $\mathcal{X} = \mathbb{R}^n$, then P_i , $i = 1, \dots, N$ are said to be an *algebraic $\bar{P}$ solution*.

Remark 3. In what follows we assume the existence of a *algebraic $\bar{P}$ matrix solution*, i.e. we assume $\mathcal{X} = \mathbb{R}^n$. Note that all statements can be modified accordingly if $\mathcal{X} \subset \mathbb{R}^n$. It is evident that any solution of the system of coupled Hamilton-Jacobi-Isaacs partial differential inequalities yields also a solution to the algebraic equations (3.24), for some Σ_i . In addition, since the integrability requirement of the solution of the Hamilton-Jacobi-Isaacs partial differential inequalities is removed in (3.24), then the set of solutions of the latter contains the set of solutions of the former system of inequalities. $\blacktriangle$

Remark 4. The system of algebraic equations (3.24), $i = 1, \dots, N$ (or (3.22) in the 2-player case) exhibits point-wise, namely for fixed $x \in \mathbb{R}^n$, a structure similar to the system of equations (3.89), which resembles what happens in the SDRE approach to optimal control problems. Therein, in fact, point-wise properties of the pairs $(F(x), g(x))$ and $(F(x), Q(x)^{1/2})$, namely controllability and observability, respectively, for all $x \in \mathbb{R}^n$, guarantees the existence of a matrix-valued function $P(x)$ that solves the SDRE and it is additionally positive definite for each fixed $x \in \mathbb{R}^n$ [83, 84]. In the framework of linear-quadratic differential games there are no such structural assumptions on the system that guarantee the existence of solutions to the coupled AREs. Therefore, clearly also the point-wise extension to the solution of (3.24) cannot be pursued in this context. $\blacktriangle$

In the following sections we show how the notion of *algebraic $\bar{P}$ matrix solution* can be exploited to determine an approximate solution – in the sense of ϵ_α -Nash equilibria discussed for linear-quadratic differential games in Section 3.3 – to Problem 4 (and hence Problem 3), without involving the solution of any partial differential equations or inequalities.

3.5 Approximate Solutions Using Individual Dynamic Extensions

In this section a method for solving the differential game defined by the dynamics (3.1), satisfying certain structural assumption, and the cost functionals (3.11) for $i = 1, \dots, N$ is presented. More specifically, the structural assumption is the following.

Assumption 4. The mappings $g_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m_i}$ are full row rank for all $x \in \mathbb{R}^n$, implying $m_i \geq n$.

The 2-player case is considered before the general N -player case.

3.5.1 2-Player Case: Approximate Solutions Using Two Individual Dynamic Extensions

Consider the differential game introduced in Problem 4, with $N = 2$. We provide a solution for a modified problem, which approximates the non-cooperative differential game introduced in Section 3.2. We firstly introduce an additional state, $\xi_i(t) \in \mathbb{R}^n$, for each individual player thus considering dynamic state feedback in place of the classical static state feedback that constitutes the Nash equilibrium control strategy. Then we show how to employ the notion of *algebraic $\bar{P}$ matrix solution* for the construction of *value functions*, one for each player, defined in the extended state-space (x, ξ_i) , that solve a system of Hamilton-Jacobi-Isaacs partial differential inequalities, instead of partial differential equations, associated to a specific *auxiliary* non-cooperative differential game. The above discussion is summarised in the following.

In what follows we consider dynamic feedback strategies of the form

$$u_i = \beta_i(x, \xi_i), \quad \dot{\xi}_i = \tau_i(x, \xi_i), \quad (3.26)$$

with $\xi_i(t) \in \mathbb{R}^\nu$, for some $\nu > 0$, $\tau_i(0, 0) = 0$, $\beta_i(0, 0) = 0$ and τ_i, β_i smooth mappings.

Let $\xi = (\xi_1, \xi_2)$. We define a set of α -admissible dynamic feedback strategies as follows.

Definition 17. A pair of dynamic feedback strategies $\mathcal{S} = \{(u_1, \dot{\xi}_1), (u_2, \dot{\xi}_2)\}$ is said to be (α -admissible) if the zero-equilibrium of the closed-loop system (3.2)-(3.26), $i = 1, 2$, is (locally) asymptotically stable ($\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of (3.2)-(3.26) around the origin).

Problem 5. Consider the system (3.2) and the cost functionals (3.3). The problem of solving the *approximate dynamic non-cooperative differential game* consists in determining a pair

of admissible *dynamic feedback strategies* $\mathcal{S} = \{S_1, S_2\}$, where the strategy S_i , $i = 1, 2$, is a dynamical system described by (3.26), with $\nu = n$, $i = 1, 2$, and non-negative functions $c_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that for any admissible pairs $\mathcal{S}_1 = \{(u_1, \tau_1), (\beta_2, \tau_2)\}$ and $\mathcal{S}_2 = \{(\beta_1, \tau_1), (u_2, \tau_2)\}$, with $u_1 \neq \beta_1$ and $u_2 \neq \beta_2$, $\hat{J}_i((x(0), \xi(0)), \beta_i, \beta_j) \leq \hat{J}_i((x(0), \xi(0)), u_i, \beta_j)$, $i = 1, 2$, $j \neq i$, where the extended cost functionals $\hat{J}_i$, $i = 1, 2$, are defined as

$$\hat{J}_i(x(0), \xi(0), u_i, u_j) \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x(t)) + \|u_i(t)\|^2 - \|u_j(t)\|^2 + c_i(x(t), \xi(t)) \right) dt. \quad (3.27)$$

Making use of the notion of algebraic $\bar{P}$ matrix solution, consider the *extended value functions*

$$V_i(x, \xi_i) = \frac{1}{2} x^\top P_i(\xi_i) x + \frac{1}{2} \|x - \xi_i\|_{R_i}^2, \quad (3.28)$$

with $\xi_i \in \mathbb{R}^n$ and $R_i = R_i^\top > 0$, $i = 1, 2$. In addition, consider the continuous matrix-valued functions $\Delta_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ defined as

$$\Delta_i(x, \xi_i) = (R_i - \Phi_i(x, \xi_i)) R_i^{-1} \Psi_i(x, \xi_i), \quad (3.29)$$

where $\Phi_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ is a continuous matrix-valued function satisfying $x^\top (P_i(x) - P_i(\xi_i)) = (x - \xi_i)^\top \Phi_i(x, \xi_i)^\top$. Furthermore, let $\Psi_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ denote the Jacobian matrix of the mapping $\frac{1}{2} P_i(\xi_i) x$ with respect to ξ_i and define

$$A_{cl}(x) \triangleq F(x) - g_1(x) g_1(x)^\top P_1(x) - g_2(x) g_2(x)^\top P_2(x). \quad (3.30)$$

Note that the vector field $A_{cl}(x)x$ describes the closed-loop nonlinear system when only the *algebraic inputs* $u_i = -g_i(x)^\top P_i(x)x$, $i = 1, 2$, are applied.

Theorem 3. Consider the system (3.2), the cost functionals (3.3) and suppose Assumption 4 holds for $i = 1, 2$. Let P_i , $i = 1, 2$, be an algebraic $\bar{P}$ matrix solution of the system (3.4). Moreover, let R_i be such that

$$\Upsilon_i \triangleq \Sigma_i - \frac{1}{2} \Delta_i^\top A_{cl} - \frac{1}{2} A_{cl}^\top \Delta_i + \frac{1}{2} \Delta_i^\top g_i g_i^\top \Delta_i > 0 \quad (3.31)$$

$$g_j^\top (P_i + P_j + \Delta_i) \Upsilon_i^{-1} (P_i + P_j + \Delta_i)^\top g_j < 4I, \quad (3.32)$$

$i = 1, 2$.⁶ Then there exist $\bar{k}_1 \geq 0$ and $\bar{k}_2 \geq 0$, and a set $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ such that the functions V_1 and V_2 in (3.28) solve the system of extended partial differential inequalities

$$\begin{aligned} \mathcal{H}\mathcal{J}_i(x, \xi) \triangleq & \frac{\partial V_i}{\partial x} f(x) + \frac{\partial V_i}{\partial \xi_i} \dot{\xi}_i - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top \\ & + \frac{1}{2} q_i(x) - \frac{1}{2} \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top - \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \leq 0, \end{aligned} \quad (3.33)$$

with $\dot{\xi}_1 = -k_1 \frac{\partial V_1}{\partial \xi_1}^\top$ and $\dot{\xi}_2 = -k_2 \frac{\partial V_2}{\partial \xi_2}^\top$, for all $k_1 > \bar{k}_1$ and $k_2 > \bar{k}_2$, and for all $(x, \xi_1, \xi_2) \in \Omega$. It follows that the dynamical systems

$$\begin{aligned} \dot{\xi}_i &= -k_i (\Psi_i(x, \xi_i)^\top x - R_i(x - \xi_i)), \\ u_i &= -g_i(x)^\top (P_i(x)x + (R_i - \Phi_i(x, \xi_i))(x - \xi_i)), \end{aligned} \quad (3.34)$$

$i = 1, 2$, are such that $\mathcal{S} = \{(u_1, \dot{\xi}_1), (u_2, \dot{\xi}_2)\}$ is admissible and solves the approximate dynamic non-cooperative differential game defined in Problem 5 with $c_i(x, \xi_i) = -2\mathcal{H}\mathcal{J}_i(x, \xi)$, $i = 1, 2$. Moreover, there exists a neighbourhood of the origin in which the set of dynamic feedback strategies $\mathcal{S}$ constitutes an ϵ_α -Nash equilibrium solution for Problem 3. $\diamond$

Proof: The first part of the claim is proved in two steps. First we show that the functions V_1 and V_2 , in (3.28) solve the system of extended partial differential inequalities defined in (3.33), provided k_1 and k_2 are sufficiently large. It follows that the dynamic control laws (3.34) minimise the extended cost functionals introduced in (3.27). In particular the functions c_i are given by the additional negativity of the inequalities (3.33), namely the *gap*, or *confidence interval*, between the inequalities, satisfied by V_1 and V_2 , and the corresponding equations. Then we show that the zero equilibrium of the closed-loop

⁶Note that $\Upsilon_i(0, 0) = \Sigma_i > 0$, $i = 1, 2$.

system

$$\begin{aligned}\dot{x} &= f(x) - g_1(x)g_1(x)^\top \frac{\partial V_1}{\partial x}^\top - g_2(x)g_2(x)^\top \frac{\partial V_2}{\partial x}^\top, \\ \dot{\xi}_1 &= -k_1 \frac{\partial V_1}{\partial \xi_1}^\top, \\ \dot{\xi}_2 &= -k_2 \frac{\partial V_2}{\partial \xi_2}^\top,\end{aligned}\tag{3.35}$$

is locally asymptotically stable, thus proving admissibility.

To begin with the partial derivatives of the functions $V_i(x, \xi_i)$ with respect to x and ξ_i are

$$\begin{aligned}\frac{\partial V_i}{\partial x} &= x^\top P_i(x) + (x - \xi_i)^\top (R_i - \Phi_i(x, \xi_i))^\top, \\ \frac{\partial V_i}{\partial \xi_i} &= x^\top \Psi_i(x, \xi_i) - (x - \xi_i)^\top R_i.\end{aligned}\tag{3.36}$$

Therefore, the i^{th} inequality of the system of partial differential inequalities (3.33), considering the partial derivatives of V_i as in (3.36), the dynamic control law as in (3.34) and recalling that the matrix-valued functions $P_i, i = 1, 2$, constitute an *algebraic $\bar{P}$ matrix solution* of the Hamilton-Jacobi-Isaacs equations (3.4) – rewritten as a *quadratic* form in x , $(x - \xi_i)$ and $(x - \xi_j)$ – yields

$$-\left[x^\top (x - \xi_i)^\top (x - \xi_j)^\top \right] \left(M_i + k_i C_i^\top C_i \right) \begin{bmatrix} x \\ x - \xi_i \\ x - \xi_j \end{bmatrix} \leq 0, \tag{3.37}$$

where the matrix-valued function $M_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{3n \times 3n}$ is given by

$$M_i = \begin{bmatrix} \Sigma_i & -\frac{1}{2}A_{cl}^\top(R_i - \Phi_i) & \frac{1}{2}(P_i^\top + P_j^\top)g_j g_j^\top(R_j - \Phi_j) \\ -\frac{1}{2}(R_i - \Phi_i)^\top A_{cl} & \frac{1}{2}(R_i - \Phi_i)^\top g_i g_i^\top(R_i - \Phi_i) & \frac{1}{2}(R_i - \Phi_i)^\top g_j g_j^\top(R_j - \Phi_j) \\ \frac{1}{2}(R_j - \Phi_j)^\top g_j g_j^\top(P_i + P_j) & \frac{1}{2}(R_j - \Phi_j)^\top g_j g_j^\top(R_i - \Phi_i) & (R_j - \Phi_j)^\top g_j g_j^\top(R_j - \Phi_j) \end{bmatrix}, \tag{3.38}$$

whereas the matrix-valued function $C_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times 3n}$ is defined as

$$C_i(x, \xi_i) = \begin{bmatrix} \Psi_i(x, \xi_i) & -R_i & 0 \end{bmatrix}. \quad (3.39)$$

Note that, since R_i is positive definite, the matrix $C_i(x, \xi_i)$ has constant rank, equal to n , for all $\xi_i \in \mathbb{R}^n$, hence the kernel of $C_i(x, \xi_i)$ has dimension $2n$.

In particular, consider the matrix

$$Z_i = \begin{bmatrix} I & 0 \\ R_i^{-1} \Psi_i(x, \xi_i) & 0 \\ 0 & I \end{bmatrix} \quad (3.40)$$

and note that the columns of the matrix $Z_i(x, \xi_i)$ span the kernel of $C_i(x, \xi_i)$ for all $\xi_i \in \mathbb{R}^n$. Following [85] we require $Z_i^\top M_i Z_i$ to be positive definite. Note that

$$Z_i^\top M_i Z_i = \frac{1}{2} \begin{bmatrix} 2\Upsilon_i & \Gamma_{i1} \\ \Gamma_{i1}^\top & \Gamma_{i2} \end{bmatrix}, \quad (3.41)$$

where $\Gamma_{i1} = (P_i + P_j + \Delta_i)g_j g_j^\top (R_j - \Phi_j)$ and $\Gamma_{i2} = 2(R_j - \Phi_j)^\top g_j g_j^\top (R_j - \Phi_j)$. The first part of the claim is then proved noting that – considering Assumption 4 and the Schur complement of the element Υ_i in the matrix (3.41) – positive definiteness of the matrix $Z_i^\top M_i Z_i$ is equivalent to the conditions (3.31)-(3.32).

To conclude the proof we need to show that the closed-loop system (3.35) has an asymptotically stable equilibrium point at the origin. Firstly, note that there exist a neighbourhood $\mathcal{W}_i$ of the origin such that the function V_i as in (3.28) is positive definite in $\mathcal{W}_i$. This can be seen recalling that, by definition P_i , is *tangent* at the origin to $\bar{P}_i > 0$. Let $W(x, \xi_1, \xi_2) = V_1(x, \xi_1) + V_2(x, \xi_2)$ be a candidate Lyapunov function. The time derivative of the function W along the trajectories of the closed-loop system (3.35) is

$$\dot{W} = \dot{V}_1 + \dot{V}_2 \leq -\frac{1}{2}q_1(x) - \frac{1}{2}q_2(x), \quad (3.42)$$

for all $x \neq 0$ in a neighbourhood of the zero equilibrium. Hence, by Assumption 2,

and exploiting LaSalle's invariance principle, $\lim_{t \rightarrow \infty} x(t) = 0$. Moreover, by (3.42), all the trajectories of the system (3.35) belong to the compact set $\{(x, \xi_1, \xi_2) : W(x, \xi_1, \xi_2) \leq W(x(0), \xi_1(0), \xi_2(0))\}$, hence are bounded. Finally, noting that the zero equilibrium of the systems $\dot{\xi}_i = -k_i \frac{\partial V_i}{\partial \xi_i}(0, \xi_i)^\top$, $i = 1, 2$, are globally asymptotically stable, we prove asymptotic stability of $(x, \xi_1, \xi_2) = (0, 0, 0)$ and detectability for the system (3.35), by standard arguments on interconnected systems. For any α -admissible set of control strategies the integral $\int_0^\infty c_i(x, \xi) dt$, $i = 1, 2$, is bounded in a neighbourhood $\bar{\Omega}$, which may depend on α , of the origin. Then, (3.33) and the results of Proposition 2 imply that in $\bar{\Omega}$ the set of strategies $\mathcal{S}$ is an ϵ_α -Nash equilibrium for Problem 3. $\square$

Remark 5. From the arguments employed in the second part of the above proof it is evident that Assumption 2 may be relaxed by assuming instead *zero-state detectability* of the system (3.35) with respect to the output $y = q_1(x) + q_2(x)$. $\blacktriangle$

Remark 6. Suppose that the matrices $\bar{\Sigma}_i$, $i = 1, 2$, are positive definite. Then there exist a non-empty $\hat{\Omega}_1 \subseteq \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ and matrices R_i such that the conditions (3.31) are satisfied for all $(x, \xi_1, \xi_2) \in \hat{\Omega}_1$. In fact note that $R_i = \Phi_i(0, 0)$ yields $\Upsilon_i(0, 0) = \bar{\Sigma}_i > 0$ and the existence of the set $\hat{\Omega}_1$ can be concluded by continuity of the functions in the left-hand side of (3.31). $\blacktriangle$

Remark 7. Suppose that the matrices $\bar{\Sigma}_i$, $i = 1, 2$, are positive definite and that

$$B_j^\top (\bar{P}_i + \bar{P}_j) \bar{\Sigma}_i^{-1} (\bar{P}_i + \bar{P}_j) B_j < 4I, \quad (3.43)$$

$i = 1, 2, j = 1, 2, i \neq j$. Then there exist a non-empty $\hat{\Omega}_2 \subseteq \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ and matrices R_i such that the conditions (3.32) are satisfied for all $(x, \xi_1, \xi_2) \in \hat{\Omega}_2$. $\blacktriangle$

The structural properties in Assumption 4, $i = 1, 2$ can be removed, provided certain conditions are satisfied, and it can be guaranteed that the inequalities (3.32) are satisfied, at least locally around the origin of the extended state-space as detailed in the following corollary.

Corollary 1. Consider the system (3.2) and the cost functionals (3.3). Suppose that $x = 0$ is a locally exponentially stable equilibrium of the system (3.2) and $g_1(0) = 0$ and $g_2(0) =$

0. Let $P_i, i = 1, 2$, be an algebraic $\bar{P}$ matrix solution of the system (3.4) with $\bar{\Sigma}_i > 0$. Let $R_i = \Phi_i(0, 0)$. Then $M_i \geq 0$ at least in a neighbourhood of the origin. In addition there exist $\bar{k}_i \geq 0, i = 1, 2$, and a set Ω such that the dynamical systems (3.34) are dynamic strategies that solve Problem 5 for all $k_i \geq \bar{k}_i, i = 1, 2$, and all $(x, \xi_1, \xi_2) \in \Omega$. $\diamond$

Proof: The claim is proved by showing that, with these assumptions, the conditions (3.31)-(3.32) are satisfied, for each player, at least locally. Since $\bar{\Sigma}_i > 0$, by Remark 6, the conditions (3.31) hold for all $(x, \xi_1, \xi_2) \in \hat{\Omega}_3$, where $\hat{\Omega}_3$ is a neighbourhood of the origin. Finally, note that, by assumption, $g_i(0) = 0$ for $i = 1, 2$, hence the conditions (3.43) are trivially satisfied. $\square$

This section is concluded by showing an interesting result of the dynamic strategies (3.34). Consider a given set of admissible strategies $\hat{\mathcal{S}}_i = \{(\hat{u}_i, \hat{\xi}_1), u_j^*\}$, where $\mathcal{S}^* = \{u_1^*, u_2^*\}$ denotes the Nash equilibrium solution of the differential game and let $\hat{x}(t)$ and $\hat{\xi}(t)$ denote the trajectory of the extended state. The integral of the function $c_i(\hat{x}(t), \hat{\xi}(t))$ provides an upper bound for the excess of the cost $J_i(\bar{u}_i, u_j^*)$ with respect to the Nash equilibrium outcome $J_i(u_i^*, u_j^*)$. $\blacktriangle$

3.5.2 N -Player Case: Approximate Solutions Using N Individual Dynamic Extensions

In this subsection we consider the general case in which the differential game has N participants. Suppose that Assumption 4 is satisfied for $i = 1, \dots, N$. Similarly, the definitions introduced in the 2-player differential game can be straight-forwardly extended to the N player case. For instance, the matrix valued function (3.30) becomes $A_{cl}(x) = F(x) - \sum_{i=1}^N g_i(x)g_i(x)^\top P_i(x)$, $\xi = (\xi_1, \dots, \xi_N)$ and α -admissible dynamic feedback strategies are defined as in the following statement.

Definition 18. The set of dynamic feedback strategies $\mathcal{S} = \{(u_1, \dot{\xi}_1), \dots, (u_N, \dot{\xi}_N)\}$ is said to be admissible (α -admissible) if the zero-equilibrium of the closed-loop system (3.1)-(3.26), $i = 1, \dots, N$, is (locally) asymptotically stable ($\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of (3.1)-(3.26) around the origin).

Problem 6. Consider the system (3.1) and the cost functionals (3.11), with $i = 1, \dots, N$. The problem of solving the *approximate dynamic non-cooperative differential game* consists in determining a set of admissible *dynamic feedback strategies* $\mathcal{S} = \{S_1, \dots, S_N\}$, where the strategy S_i , $i = 1, \dots, N$, is a dynamical system described by equations of the form (3.26), with $\nu = n$, $i = 1, \dots, N$, and non-negative functions $c_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that for any admissible set of dynamic strategies $\mathcal{S}_i = \{(\beta_1, \tau_1), \dots, (u_i, \tau_i), \dots, (\beta_N, \tau_N)\}$, with $u_i \neq \beta_i$, the following holds for $i = 1, \dots, N$, $\hat{J}_i((x(0), \xi(0)), \beta_1, \dots, \beta_i, \dots, \beta_N) \leq \hat{J}_i((x(0), \xi(0)), \beta_1, \dots, u_i, \dots, \beta_N)$, where the extended cost functionals $\hat{J}_i$, $i = 1, \dots, N$, are given by

$$\hat{J}_i \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x) + \|u_i\|^2 - \sum_{j=1, j \neq i}^N \|u_j\|^2 + c_i(x, \xi) \right) dt. \quad (3.44)$$

The following statement provides a solution to the approximate dynamic non-cooperative differential game for N players, thus extending the result of Theorem 3.

Theorem 4. Consider the system (3.1) and the cost functionals (3.11), where $i = 1, \dots, N$. Suppose that Assumption 4 holds. Let P_i , $i = 1, \dots, N$, be solutions of the system (3.24) and (3.89). Let R_i be such that

$$\Upsilon_i \triangleq \Sigma_i - \frac{1}{2} \Delta_i^\top A_{cl} - \frac{1}{2} A_{cl}^\top \Delta_i + \frac{1}{2} \Delta_i^\top g_i g_i^\top \Delta_i > 0 \quad (3.45)$$

$$\Pi_i \Upsilon_i^{-1} \Pi_i^\top \leq 4 \text{blockdiag}\{I\}, \quad (3.46)$$

with

$$\Pi_i \triangleq \begin{bmatrix} g_1(x)^\top (P_1(x) + P_i(x) + \Delta_i) \\ \vdots \\ g_{i-1}(x)^\top (P_{i-1}(x) + P_i(x) + \Delta_i) \\ g_{i+1}(x)^\top (P_{i+1}(x) + P_i(x) + \Delta_i) \\ \vdots \\ g_N(x)^\top (P_N(x) + P_i(x) + \Delta_i) \end{bmatrix}, \quad (3.47)$$

and $i = 1, \dots, N$.

Then there exist $\bar{k}_i \geq 0$, $i = 1, \dots, N$, and a set $\Omega \subseteq \mathbb{R}^n \times \{\mathbb{R}^n \times \dots \times \mathbb{R}^n\}$ such that the functions V_i , $i = 1, \dots, N$, as in (3.28) solve the system of extended partial differential

inequalities

$$\begin{aligned} \mathcal{H}\mathcal{J}_i \triangleq & \frac{\partial V_i}{\partial x} f(x) + \frac{\partial V_i}{\partial \xi_i} \dot{\xi}_i - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) \\ & - \frac{1}{2} \sum_{j=1, j \neq i}^N \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top - \sum_{j=1, j \neq i}^N \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \leq 0, \end{aligned} \quad (3.48)$$

with $\dot{\xi}_i = -k_i \frac{\partial V_i}{\partial \xi_i}^\top$, $i = 1, \dots, N$, for all $k_i > \bar{k}_i$ and for all $(x, \xi) \in \Omega$. Suppose additionally that

$$\sum_{i=1}^N \left(\frac{N-2}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top - c_i(x, \xi) \right) \leq 0, \quad (3.49)$$

where $c_i(x, \xi) = -2\mathcal{H}\mathcal{J}_i(x, \xi)$, for all $(x, \xi_i) \in \mathbb{R}^n \times \mathbb{R}^n$, $i = 1, \dots, N$. Then, the dynamical systems, for $i = 1, \dots, N$,

$$\begin{aligned} \dot{\xi}_i &= -k_i (\Psi_i(x, \xi_i)^\top x - R_i(x - \xi_i)), \\ u_i &= -g_i(x)^\top (P_i(x)x + (R_i - \Phi_i(x, \xi_i))(x - \xi_i)), \end{aligned} \quad (3.50)$$

are such that $\mathcal{S} = \{(u_1, \dot{\xi}_1), \dots, (u_N, \dot{\xi}_N)\}$ is admissible and solves Problem 6. Moreover, there exists a neighbourhood of the origin in which the set of dynamic feedback strategies $\mathcal{S}$ constitutes an ϵ_α -Nash equilibrium solution for Problem 4. $\diamond$

Proof: The first part of the claim is proved following the same steps as those in the proof of Theorem 3. To prove stability of the origin, i.e. admissibility of the dynamic strategies, for the closed-loop extended system

$$\dot{x} = f(x) - \sum_{i=1}^N g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top, \quad \dot{\xi}_i = -k_i \frac{\partial V_i}{\partial \xi_i}^\top, \quad i = 1, \dots, N, \quad (3.51)$$

let $W(x, \xi_1, \dots, \xi_N) = \sum_{i=1}^N V_i(x, \xi_i)$ be a candidate Lyapunov function. The time derivative of W along the trajectories of the system (3.51) is, by (3.48), $\dot{W}(x) = \sum_{i=1}^N (-1/2 q_i + (N-2)/2 (\partial V_i / \partial x) g_i g_i^\top (\partial V_i / \partial x)^\top - c_i(x, \xi)) \leq -1/2 \sum_{i=1}^N q_i(x)$, where the last inequality follows from the condition (3.49). Hence the claim is proved by Assumption 2 and LaSalle's invariance principle. The last part of the claim is proved as in the proof of Theorem 3. $\square$

Remark 8. The conditions (3.45) are identical to the conditions (3.31), namely they are independent from the number of players. In fact, the condition (3.45) is related to the individual optimality of each single player in the absence of competitors, see also [63] for more details. Finally note that the conditions (3.46) represent the natural extension of the conditions (3.32) to the N -player case. ▲

Remark 9. The condition (3.49) describes the fact that the function V_i must enforce a desired *margin* of negativity in the extended partial differential inequalities (3.48) to guarantee asymptotic stability of the closed-loop system (3.51). ▲

3.6 Approximate Solutions Using a Shared Dynamic Extensions

In this section we relax the structural assumption on the mappings $g_i, i = 1, \dots, N$, namely Assumption 4. The result is achieved by using a *single* dynamic extension that is *shared* among the players and introducing additional assumptions on the *algebraic* $\bar{P}$ *matrix solution* in the N -player case. The 2-player case is considered first.

3.6.1 2-Player Case: Approximate Solutions Using a Shared Dynamic Extension

Consider the differential game defined in Problem 3. A modified problem which approximates this differential game can be formulated by introducing a dynamic extension that is *shared* between the two players. With the exception that a *shared* dynamic extension is used in place of *individual* dynamic extensions, the modified problem is similar to Problem 5.

We now consider dynamic feedback strategies of the form

$$u_i = \beta_i(x, \xi), \quad \dot{\xi} = \tau(x, \xi), \quad (3.52)$$

with $\xi_i(t) \in \mathbb{R}^\nu$, for some $\nu > 0$, $\tau(0, 0) = 0$, $\beta_i(0, 0) = 0$ and τ, β_i smooth mappings.

Definition 19. The set of dynamic feedback strategies⁷ $\mathcal{S} = \{u_1, u_2, \dot{\xi}\}$ is said to be admissible (α -admissible) if the zero-equilibrium of the closed-loop system (3.2)-(3.52), $i = 1, 2$, is (locally) asymptotically stable ($\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of (3.1)-(3.26) around the origin).

Problem 7. Consider the system (3.2) and the cost functionals (3.3). The problem of solving the *approximate dynamic non-cooperative differential game* consists in determining a pair of admissible *dynamic feedback strategies* $\mathcal{S} = \{S_1, S_2\}$, where the strategy S_i , $i = 1, 2$, is a dynamical system described by equations of the form (3.52) with $\nu = n$, $i = 1, 2$, and non-negative functions $c_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that for any admissible set of strategies $\mathcal{S}_i = \{u_i, \beta_j, \alpha\}$, with $u_i \neq \beta_i$, $\hat{J}_i((x(0), \xi(0)), \beta_i, \beta_j) \leq \hat{J}_i((x(0), \xi(0)), u_i, \beta_j)$, where the extended cost functionals $\hat{J}_i$ are defined as

$$\hat{J}_i \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x(t)) + \|u_i(t)\|^2 - \|u_j(t)\|^2 + c_i(x(t), \xi(t)) \right) dt. \quad (3.53)$$

Theorem 5. Consider the system (3.2) and the cost functionals (3.3). Suppose Assumption 2 is satisfied and let P_i , $i = 1, 2$, be an *algebraic $\bar{P}$ matrix solution* of the equations (3.4) satisfying $\bar{\Sigma}_i > 0$, $i = 1, 2$. Let R_1 and R_2 be positive definite and symmetric matrices and such that

$$R_i (R_1 + R_2) + (R_1 + R_2) R_i > 0, \quad (3.54)$$

$i = 1, 2$. Then there exist a constant $\bar{k} \geq 0$ and a set $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^n$ such that the functions

$$V_i(x, \xi) = \frac{1}{2} x^\top P_i(\xi) x + \frac{1}{2} \|x - \xi\|_{R_i}^2, \quad (3.55)$$

$i = 1, 2$, solve the system of partial differential inequalities

$$\begin{aligned} \mathcal{H}\mathcal{J}_i^s \triangleq & \frac{\partial V_i}{\partial x} f(x) + \frac{\partial V_i}{\partial \xi} \dot{\xi} - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) - \frac{1}{2} \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \\ & - \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \leq 0, \end{aligned} \quad (3.56)$$

⁷The dynamic strategy of player i is given by the tuple $(u_i, \dot{\xi})$. Since the dynamic extension is shared by all players, with some abuse of notation, $\{u_1, u_2, \dot{\xi}\}$ is used to denote that both player 1 and player 2 adhere to the strategies $(u_1, \dot{\xi})$ and $(u_2, \dot{\xi})$, respectively.

$i = 1, 2, j = 1, 2, i \neq j$, with $\dot{\xi} = -k \left(\frac{\partial V_1}{\partial \xi} + \frac{\partial V_2}{\partial \xi} \right)^\top$, for all $k > \bar{k}$ and for all $(x, \xi) \in \Omega$.

Hence the dynamical system

$$\begin{aligned}\dot{\xi} &= -k \sum_{i=1}^2 \left(\Psi_i(x, \xi)^\top x - R_i(x - \xi) \right), \\ u_i &= -g_i(x)^\top (P_i(x)x + (R_i - \Phi_i(x, \xi))(x - \xi)),\end{aligned}\tag{3.57}$$

$i = 1, 2$, is such that $\mathcal{S}$ is admissible and solves Problem 7 with $c_i(x, \xi) = -2\mathcal{H}\mathcal{J}_i^s(x, \xi)$, $i = 1, 2$. Moreover, there exists a neighbourhood of the origin in which the set of dynamic feedback strategies $\mathcal{S}$ constitutes an ϵ_α -Nash equilibrium solution for Problem 3. $\diamond$

Proof: The claim is proved in two steps following the same arguments as those exploited in the proof of Theorem 3. Therefore, we firstly show that the functions V_i , $i = 1, 2$, in (3.55) solve the system of extended partial differential inequalities defined in (3.56), provided k is sufficiently large, which implies that the dynamic control laws (3.57) minimise the extended cost functionals introduced in (3.53). In particular the functions c_i are given by the negativity of the inequalities (3.56), namely the *gap* between the inequalities, satisfied by the V_i , $i = 1, 2$, and the corresponding equalities. Then we show that the zero equilibrium of the closed-loop system

$$\begin{aligned}\dot{x} &= f(x) - g_1(x)g_1(x)^\top \frac{\partial V_1}{\partial x}^\top - g_2(x)g_2(x)^\top \frac{\partial V_2}{\partial x}^\top, \\ \dot{\xi} &= -k \left(\frac{\partial V_1}{\partial \xi} + \frac{\partial V_2}{\partial \xi} \right)^\top,\end{aligned}\tag{3.58}$$

is locally asymptotically stable, thus proving admissibility of the dynamic strategies. Mimicking the proof of Theorem 3, the i^{th} inequality of the system of partial differential inequalities (3.56), the dynamic control law as in (3.57) and recalling that the mappings P_i , $i = 1, 2$, is an *algebraic $\bar{P}$ matrix solution* of the Hamilton-Jacobi-Isaacs equations (3.4) – rewritten as a *quadratic* form in x and $(x - \xi)$ – yields

$$- \left[x^\top (x - \xi)^\top \right] (M_i + kD_i) \begin{bmatrix} x \\ x - \xi \end{bmatrix},\tag{3.59}$$

where $M_i : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{2n \times 2n}$ is a the matrix-valued function given by

$$M_i = \begin{bmatrix} \Sigma_i & \Gamma_{12} \\ \Gamma_{12}^\top & \Gamma_{22} \end{bmatrix}, \quad (3.60)$$

with $\Gamma_{12} = -\frac{1}{2}A_{cl}(x)^\top(R_i - \Phi_i) + \frac{1}{2}(P_1 + P_2)^\top g_j g_j^\top(R_j - \Phi_j)$ and $\Gamma_{22} = \frac{1}{2} \sum_{l=1}^2 (R_l - \Phi_l)^\top g_l g_l^\top(R_l - \Phi_l) + (R_i - \Phi_i)^\top g_j g_j^\top(R_j - \Phi_j)$, and

$$D_i = \frac{1}{2} \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}^\top & \Lambda_{22} \end{bmatrix}, \quad (3.61)$$

with $\Lambda_{11} = \Psi_i(\Psi_1 + \Psi_2)^\top + (\Psi_1 + \Psi_2)\Psi_i^\top$, $\Lambda_{22} = R_i(R_1 + R_2) + (R_1 + R_2)R_i$ and $\Lambda_{12} = -\Psi_i(R_1 + R_2) - (\Psi_1 + \Psi_2)R_i$. The inequality (3.59) is derived by noting that the partial derivatives of the functions V_i , $i = 1, 2$, as in (3.55) are obtained similarly to (3.36) with $\xi_1 = \xi_2 = \xi$ and the matrix-valued function $\Phi_i(x, \xi)$ is such that $x^\top (P_i(x) - P_i(\xi)) = (x - \xi)^\top \Phi_i(x, \xi)$.

Since $\Psi_i(0, \xi) = 0$, $i = 1, 2$, the inequality (3.54) ensure the existence of a neighbourhood $\mathcal{W}_i$ of the origin of $\mathbb{R}^n \times \mathbb{R}^n$ in which the matrix D_i is positive semidefinite. Moreover, since the columns of the matrix $Z = [I, 0]^\top$ span the kernel of the matrix-valued function D_i evaluated at $\xi = 0$ and $Z^\top M_i \Big|_{(0,0)} Z = \bar{\Sigma}_i$, then, by [85], there exists a non-empty open set Ω containing the origin and $\bar{k} \geq 0$ such that the inequalities (3.56), $i = 1, 2$, are satisfied for all $k > \bar{k}$ and for all $(x, \xi) \in \Omega$.

Finally, (local) asymptotic stability of the zero-equilibrium of the closed-loop system (3.58) can be shown exactly as in the proof of Theorem 3, namely employing $W(x, \xi) = V_1(x, \xi) + V_2(x, \xi)$ as a Lyapunov function and relying on Assumption 2. The last part of the claim is proved as in the proof of Theorem 3. $\square$

Remark 10. Alternative conditions, which are stronger than those given in (3.54), for ensuring the dynamic strategies (3.57), $i = 1, 2$, are admissible and solve Problem 7 are

$$(\Psi_i \Psi_j^\top + \Psi_j \Psi_i^\top) > 0, \quad (3.62)$$

and

$$(R_i R_j + R_j R_i) - \frac{1}{2} \Xi^\top (\Psi_i \Psi_j^\top + \Psi_j \Psi_i^\top)^{-1} \Xi \geq 0, \quad (3.63)$$

with $\Xi = (\Psi_i R_j + \Psi_j R_i)$. These ensure that the cross terms $\frac{\partial V_i}{\partial \xi} \frac{\partial V_j}{\partial \xi}^\top$, $i = 1, 2$, $j = 1, 2$, $j \neq i$ are non-negative. Note that the condition $\Lambda_{11} > 0$ in (3.61) is implied by (3.62) and $D_i > 0$ is implied by (3.63), in the case $R_i = \bar{P}_i$, $i = 1, 2$. The matrix

$$Z_i = \begin{pmatrix} I \\ R_i^{-1} \Psi_i^\top \end{pmatrix}$$

spans the kernel of $\hat{D}_i$, where $\hat{D}_i$ satisfies $\frac{\partial V_i}{\partial \xi} \frac{\partial V_i}{\partial \xi}^\top = \left(x^\top (x - \xi)^\top \right) \hat{D}_i \begin{pmatrix} x \\ x - \xi \end{pmatrix}$, $i = 1, 2$.

It follows from [85] that provided $\Sigma_i > 0$ and

$$\Sigma_i + \Psi_i R_i^{-1} \Gamma_{12}^\top + \Gamma_{12} R_i^{-1} \Psi_i^\top + \Psi_i R_i^{-1} \Gamma_{22} R_i^{-1} \Psi_i^\top > 0,$$

there exists $\bar{k}$ such that for all $k > \bar{k}$, $M_i + k \hat{D}_i > 0$ and the cross terms $\frac{\partial V_i}{\partial \xi} \frac{\partial V_j}{\partial \xi}^\top$ provide additional negativity to the Hamilton-Jacobi-Isaacs partial differential inequalities (3.56) if conditions (3.62) and (3.63) are satisfied for $i = 1, 2$, $j = 1, 2$ and $j \neq i$.

Asymptotic stability of the origin can be proved along the same lines as in the proof of Theorem 5. ▲

Remark 11. Although the conditions for Theorem 5 to hold, namely (3.54), are simpler than those for Theorem 3 or the alternative conditions provided in Remark 10, the proof of Theorem 5 relies on local properties of the matrices D_i and M_i in (3.59), $i = 1, 2$. Therefore, for some problems, in particular if the extended state is such that it is not in one or both of the neighbourhoods $\mathcal{W}_i$, $i = 1, 2$, the alternative conditions of Remark 10, or Theorem 3, when Assumption 4 is satisfied by all the players, may be more useful, despite the more complicated conditions that must be satisfied by the *algebraic* $\bar{P}$ matrix solution. ▲

3.6.2 N -Player Case: Approximate Solutions Using a Shared Dynamic Extension

Consider now the case in which there are N players. The modified problem is a natural extension of Problem 7. The following theorem extends Theorem 5 to the case in which an approximate solution of a nonzero-sum differential game with N players is sought.

Theorem 6. Consider the system (3.1) and the cost functionals (3.11), with $i = 1, \dots, N$ and suppose these satisfy Assumption 3. Let P_i , $i = 1, \dots, N$ be an algebraic $\bar{P}$ matrix solution, satisfying $\bar{\Sigma}_i > 0$, $i = 1, \dots, N$. Suppose $R_i = R_i^\top > 0$ is such that

$$R_i \left(\sum_{l=1, l \neq i}^N R_l \right) + \left(\sum_{l=1, l \neq i}^N R_l \right) R_i > 0, \quad (3.64)$$

for $i = 1, \dots, N$. Then there exists a constant $\bar{k} > 0$ and a set $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^n$ such that the functions V_i , $i = 1, \dots, N$, as in (3.55) solve the system of extended partial differential inequalities

$$\begin{aligned} \mathcal{HJ}_i^s \triangleq & \frac{\partial V_i}{\partial x} f(x) + \frac{\partial V_i}{\partial \xi} \dot{\xi} - \frac{1}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top + \frac{1}{2} q_i(x) \\ & - \sum_{j=1, j \neq i}^N \frac{1}{2} \frac{\partial V_j}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top - \sum_{j=1, j \neq i}^N \frac{\partial V_i}{\partial x} g_j(x) g_j(x)^\top \frac{\partial V_j}{\partial x}^\top \leq 0, \end{aligned} \quad (3.65)$$

$i = 1, \dots, N$ with $\dot{\xi} = -k \sum_{i=1}^N \frac{\partial V_i}{\partial \xi}$, for all $k > \bar{k}$ and for all $(x, \xi) \in \Omega$. Suppose additionally that

$$\sum_{i=1}^N \left(\frac{N-2}{2} \frac{\partial V_i}{\partial x} g_i(x) g_i(x)^\top \frac{\partial V_i}{\partial x}^\top - c_i(x, \xi) \right) \leq 0, \quad (3.66)$$

with $c_i(x, \xi) = -2\mathcal{HJ}_i^s(x, \xi)$, $i = 1, \dots, N$. Then the dynamical system

$$\begin{aligned} \dot{\xi} &= -k \sum_{i=1}^N \left(\Psi_i(x, \xi)^\top x - R_i(x - \xi) \right), \\ u_i &= -g_i(x)^\top (P_i(x)x + (R_i - \Phi_i(x, \xi))(x - \xi)), \end{aligned} \quad (3.67)$$

$i = 1, \dots, N$, is such that $\mathcal{S} = \{u_1, \dots, u_N, \dot{\xi}\}$ is admissible and solves the approximate dynamic non-cooperative differential game defined by the dynamics (3.1) and the cost

functionals (3.11). Moreover, there exists a neighbourhood of the origin in which the set of dynamic feedback strategies $\mathcal{S}$ constitutes an ϵ_α -Nash equilibrium solution for Problem 4. $\diamond$

Proof: As in the 2-player case the inequalities (3.65) can be written in quadratic form in x and $x - \xi$ as in (3.59) with M_i and D_i given by (3.60) and (3.61) with

$$\Gamma_{12} = -\frac{1}{2}A_{cl}(x)^\top (R_i - \Phi_i(x, \xi)) + \frac{1}{2} \sum_{j=1, j \neq i}^N (P_i(x) + P_j(x))^\top g_j(x)g_j(x)^\top (R_j - \Phi_j(x, \xi)),$$

$$\begin{aligned} \Gamma_{22} = & \frac{1}{2} \sum_{l=1}^N (R_l - \Phi_l(x, \xi))^\top g_l(x)g_l(x)^\top (R_l - \Phi_l(x, \xi)) \\ & + \sum_{j=1, j \neq i}^N (R_i - \Phi_i(x, \xi))^\top g_j(x)g_j(x)^\top (R_j - \Phi_j(x, \xi)) \end{aligned}$$

and $\Lambda_{11} = \Psi_i \sum_{l=1}^N \Psi_l^\top + \sum_{l=1}^N \Psi_l^\top \Psi_i$, $\Lambda_{22} = R_i \sum_{l=1}^N R_l + \sum_{l=1}^N R_l R_i$, $\Lambda_{12} = -\Psi_i \sum_{l=1}^N R_l - \sum_{l=1}^N \Psi_l R_i$, with $i = 1, \dots, N$. Similarly to the 2-player case, since $\Psi_i(0, \xi) = 0$, $i = 1, \dots, N$, the inequality (3.64) ensure the existence of a neighbourhood $\mathcal{W}_i$ of the origin where the matrix D_i is positive semidefinite and the columns of the matrix $Z = [I, I]^\top$ span the kernel of the matrix-valued function D_i evaluated at $\xi = 0$. Since the matrices R_i , $i = 1, \dots, N$, satisfy (3.64) $Z^\top M_i \Big|_{(0,0)} Z = \bar{\Sigma}_i$ and by the same arguments used in the 2-player case it can be concluded that there exists a non-empty open set Ω containing the origin and $\bar{k} \geq 0$ such that the inequality (3.56), $i = 1, \dots, N$, is satisfied for all $k > \bar{k}$ and for all $(x, \xi) \in \Omega$.

Finally, consider the candidate Lyapunov function $W(x, \xi) = \sum_{i=1}^N V_i(x, \xi)$ and note that inequality (3.66) implies that $\dot{W} \leq -\frac{1}{2} \sum_{i=1}^N q_i(x)$. Hence local asymptotic stability of the origin of the closed-loop system can be shown as in the proof of Theorem 5. The last part of the claim is proved as in the proof of Theorem 3. $\square$

Remark 12. Similarly to Remark 10, it is possible to derive alternative conditions to (3.64) ensuring that the cross terms $\frac{\partial V_i}{\partial \xi} \frac{\partial V_j}{\partial \xi}^\top$ are non-negative. $\blacktriangle$

3.7 Numerical Examples

Two numerical examples illustrating the methods for solving nonzero-sum differential games approximately using dynamic state feedback are presented. The aim of this section is to illustrate how the theory presented in this chapter can be used to obtain approximate solutions to differential games and thus provide a *qualitative* insight to the theory. An academic example where the Nash equilibrium solution is known and used as a reference for comparison between the dynamic approximations and the linear-quadratic approximation of the problem is considered first, followed by a second example involving a *competitive Lotka-Volterra model* with two competing species.

3.7.1 Academic Example

Consider the input-affine dynamical system

$$\begin{aligned} \dot{x} = & \begin{bmatrix} a_1 x_1 + 2a_2 \frac{x_1 x_2^2}{1+x_2^2} \\ a_2 x_2 \end{bmatrix} + \begin{bmatrix} m_1(1+x_2^2) & 2m_2 \frac{x_1 x_2}{1+x_2^2} \\ 0 & m_2 \end{bmatrix} u_1 \\ & + \begin{bmatrix} n_1(1+x_2^2) & 2n_2 \frac{x_1 x_2}{1+x_2^2} \\ 0 & n_2 \end{bmatrix} u_2, \end{aligned} \quad (3.68)$$

where a_1, a_2, m_1, m_2, n_1 and n_2 are constant parameters, and consider the two-player nonzero-sum differential game, where the first and the second players seek to minimise the cost functionals (3.3), $i = 1, 2$, with the running costs

$$\begin{aligned} q_1(x) &= \alpha_1 \frac{x_1^2}{(1+x_2^2)^2} + \alpha_2 x_2^2, \\ q_2(x) &= \beta_1 \frac{x_1^2}{(1+x_2^2)^2} + \beta_2 x_2^2, \end{aligned} \quad (3.69)$$

respectively, where $\alpha_1 > 0, \alpha_2 > 0, \beta_1 > 0$ and $\beta_2 > 0$ are scalar parameters.

Suppose the system parameters a_1 and a_2 are such that $a_1 < 0$ and $a_2 < 0$. An

algebraic $\bar{P}$ matrix solution for this nonzero-sum differential game is

$$P_1(x) = -\text{diag}\left(\hat{\alpha}_1 a_1 \left(\frac{1}{(1+x_2^2)^2} + c_1 x_1^2 + c_2 x_2^2\right), \hat{\alpha}_2 a_2 (1 + c_3 x_1^2 + c_4 x_2^2)\right), \quad (3.70)$$

and

$$P_2(x) = -\text{diag}\left(\hat{\beta}_1 a_2 \left(\frac{1}{(1+x_2^2)^2} + d_1 x_1^2 + d_2 x_2^2\right), \hat{\beta}_2 a_2 (1 + d_3 x_1^2 + d_4 x_2^2)\right), \quad (3.71)$$

with $c_i \geq 0$, $i = 1, \dots, 4$, $d_i \geq 0$, $i = 1, \dots, 4$, and with $\hat{\alpha}_i$ and $\hat{\beta}_i$, $i = 1, 2$ satisfying $\hat{\alpha}_1 a_1^2 \geq \alpha_1$, $\hat{\alpha}_2 a_2^2 \geq \alpha_2$, $\hat{\beta}_1 a_1^2 \geq \beta_1$ and $\hat{\beta}_2 a_2^2 \geq \beta_2$ such that Σ_1 and Σ_2 in (3.22) are positive definite.

The linear-quadratic approximation of the problem is given by the linear system (3.6) with

$$A = \text{diag}\{a_1, a_2\}, B_1 = \text{diag}\{m_1, m_2\}, B_2 = \text{diag}\{n_1, n_2\} \quad (3.72)$$

and the quadratic approximation of the running costs,

$$\begin{aligned} q_1^l &= x^\top \text{diag}\{\alpha_1, \alpha_2\} x, \\ q_2^l &= x^\top \text{diag}\{\beta_1, \beta_2\} x. \end{aligned} \quad (3.73)$$

Note that by performing the nonlinear change of coordinates $\hat{x}_1 = \frac{x_1}{1+x_2^2}$, $\hat{x}_2 = x_2$, the new state $\hat{x} = [\hat{x}_1, \hat{x}_2]^\top$ satisfies the linear dynamics (3.6) with the matrices A , B_1 and B_2 in (3.72) and, additionally, in these coordinates the running costs are $\hat{q}_1 = \alpha_1 \hat{x}_1^2 + \alpha_2 \hat{x}_2^2$ and $\hat{q}_2 = \beta_1 \hat{x}_1^2 + \beta_2 \hat{x}_2^2$. Thus, the nonzero-sum differential game boils down to a linear-quadratic game, the Nash equilibrium solution of which is given by the strategies (3.8) and the solution to the coupled AREs (3.9) in the new coordinates $\hat{x}$.

Suppose now that the system parameters are $a_1 = a_2 = -0.5$, $m_1 = m_2 = n_2 = 0.1$, $n_1 = 0.2$, $\alpha_1 = 32$, $\alpha_2 = 0.5$, $\beta_1 = 28$ and $\beta_2 = 5.25$. For this particular selection the solution to the coupled AREs associated with the linear-quadratic game in the coordinates $\hat{x}$ are $\hat{P}_1 = \text{diag}\{10, 0\}$ and $\hat{P}_2 = \text{diag}\{15, 5\}$, and it follows that the Nash equilibrium strategies in the original coordinates are $u_1^* = -B_1^\top \hat{P}_1 \left[\frac{x_1}{1+x_2^2}, x_2 \right]^\top$ and

$u_2^* = -B_2^\top \hat{P}_2 \left[\frac{x_1}{1+x_2^2}, x_2 \right]^\top$, whereas the strategies resulting from the linear-quadratic approximation of the problem are $u_1^l = -B_1^\top \hat{P}_1 x_1$ and $u_2^l = -B_2^\top \hat{P}_2 x_2$. Taking (3.70) and (3.71) as an *algebraic $\bar{P}$ matrix solution*, with $c_1 = c_2 = c_3 = 0$, $c_4 = 0.2$, $d_1 = d_2 = d_3 = 0$, $d_4 = 0.8$, $\hat{\alpha}_1 = 64$, $\hat{\alpha}_2 = 0.5$, $\hat{\beta}_1 = 56$ and $\hat{\beta}_2 = 10.5$, consider first the dynamic approximation in which each player is associated with an *individual* dynamic extension. Let u_1^m and u_2^m , given by (3.34), denote the dynamic feedback strategies of player 1 and player 2, respectively. Similarly, using the same *algebraic $\bar{P}$ matrix solution*, the dynamic approximation using a *shared* dynamic extension, yields the strategies u_1^s and u_2^s , given by (3.57), for player 1 and player 2, respectively. In summary, we have four different pairs of strategies at our disposal, namely the Nash equilibrium strategies $\mathcal{S}^* = \{u_1^*, u_2^*\}$, the linear strategies $\mathcal{S}^l = \{u_1^l, u_2^l\}$, and the two dynamic strategies⁸ $\mathcal{S}^m = \{u_1^m, u_2^m\}$ and $\mathcal{S}^s = \{u_1^s, u_2^s\}$. Since u_1^* and u_2^* are Nash equilibrium strategies, $J_i(u_i^*, u_j^*) \leq J_i(u_i^\bullet, u_j^*)$, where $u_i^\bullet$ denotes either u_i^l , u_i^m or u_i^s .

To compare the three approximate solutions of the differential game the quantities

$$C_1^d(x_0) = \begin{cases} \frac{J_1(u_1^d, u_2^*) - J_1(u_1^l, u_2^*)}{|J_1(u_1^*, u_2^*)|}, & \text{if } J_1(u_1^*, u_2^*) \neq 0 \end{cases}$$

$$C_2^d(x_0) = \begin{cases} \frac{J_2(u_1^*, u_2^d) - J_2(u_1^*, u_2^l)}{|J_2(u_1^*, u_2^*)|}, & \text{if } J_2(u_1^*, u_2^*) \neq 0 \end{cases}$$

are introduced, where d denotes either of the dynamic strategies, *i.e.* d denotes either m or s . Loosely speaking, C_i^d quantifies the difference between the outcomes when player i adopts u_i^d and u_i^l while assuming the opponent adheres to its Nash equilibrium strategy. Thus, $C_i^d < 0$ implies that the outcome $J_i(u_i^d, u_j^*)$ is closer to the Nash equilibrium outcome than the outcome $J_i(u_i^l, u_j^*)$, suggesting that the performance of the dynamic strategy u_i^d is *closer* to the Nash equilibrium strategy, u_i^* , of player i than the linear strategy u_i^l . The converse is true if $C_i^d > 0$.

Simulations have been run for several initial conditions in the positive orthant: 25

⁸To aid the clarity of presentation the dynamics of the extensions ξ_1 and ξ_2 , for the dynamic strategies using *individual* dynamic extensions, and the dynamics of ξ , for the dynamic strategies using a *shared* dynamic extension, have not been included in the notation for the dynamic feedback strategies u_1^m , u_2^m , u_1^s and u_2^s , and the sets of dynamic strategies $\mathcal{S}^m$ and $\mathcal{S}^s$.

initial conditions have been selected such that they *cover* uniformly a square grid. Similar considerations apply to the other quadrants. For the approximate dynamic solutions, u_i^m and u_i^s , different control parameters have been adopted, *i.e.* in (3.34), different $k_1, k_2, R_1, R_2, \xi_1^m(0)$ and $\xi_2^m(0)$ have been selected for different x_0 . Similarly, in (3.57), different k, R_1, R_2 and $\xi^s(0)$ have been selected for each x_0 .

Figures 3.1 and 3.2 show the values taken by C_1^m and C_2^m , respectively, across the grid of initial conditions $x_0 = [x_{1,0}, x_{2,0}]^\top$. Similarly, Figures 3.3 and 3.4 show C_1^s and C_2^s , respectively, corresponding to the different initial conditions. Close to the origin the linear system (3.6) with the matrices (3.72) is a close approximation of the nonlinear system (3.68). Similarly, along the line $x_2 = 0$, the system (3.68) and its linear approximation are identical, and it follows that for initial conditions, x_0 , close to this line (3.6)-(3.72) is a good approximation of (3.68). Consequently, in these regions it is expected that the linear strategies u_1^l and u_2^l are similar to the Nash equilibrium strategies u_1^* and u_2^* . The simulations verify this, as $J_1(u_1^l, u_2^*)$ and $J_1(u_1^*, u_2^*)$, and similarly $J_1(u_1^*, u_2^l)$ and $J_2(u_1^*, u_2^*)$, are identical or very close to one another in these regions. Despite the linear strategies being very close, or even identical, to the Nash equilibrium solution in these regions, the dynamic solutions yield comparable, in fact very similar, performances. From Figures 3.1, 3.2, 3.3 and 3.4 it is clear that both approximate dynamic strategies perform close to the linear strategies in these regions, as $C_1^m(x_0), C_2^m(x_0), C_1^s(x_0)$ and $C_2^s(x_0)$ are close to zero for x_0 close to the origin or close to the line $x_2 = 0$. Across the grid, the largest values taken by C_1^m, C_2^m, C_1^s and C_2^s are 0.0142, 0.0269, 0.0160, 0.0270 and these all correspond to the initial condition $x_0 = [2, 0.5]^\top$. Thus, this result demonstrates that even in this region the dynamic strategies u_i^m and $u_i^s, i = 1, 2$, perform relatively well. Moreover, the figures show that relatively far from the origin and the line $x_2 = 0$ both $C_i^m < 0$ and $C_i^s < 0$, indicating that the approximate dynamic strategies result in performances closer to the Nash equilibrium outcomes than the linear strategies in this region.

Figure 3.5 shows the trajectories of the state x for a small selection of x_0 , when the players adopt the strategy pairs S^* (solid, gray line), S^l (solid, black line), S^m (dashed line) and S^s (dotted line). The solid circles denote the initial conditions x_0 . Interestingly,

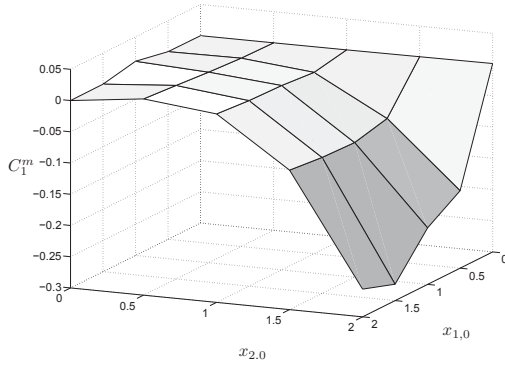


Figure 3.1: The values of C_1^m corresponding to the different initial conditions, x_0 .

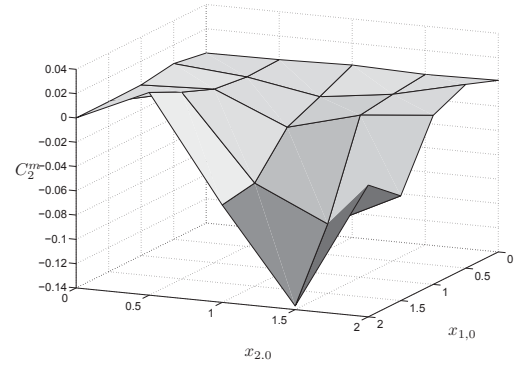


Figure 3.2: The values of C_2^m corresponding to the different initial conditions, x_0 .

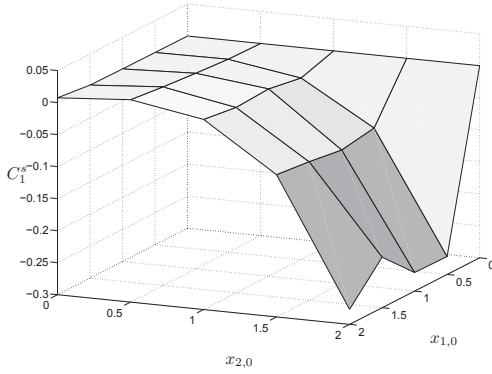


Figure 3.3: The values of C_1^s corresponding to the different initial conditions, x_0 .

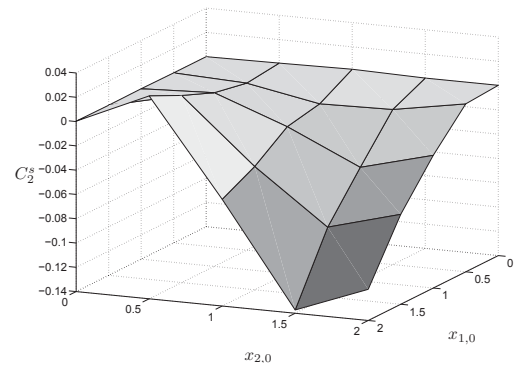


Figure 3.4: The values of C_2^s corresponding to the different initial conditions, x_0 .

for the selection of initial conditions shown, the trajectories corresponding to both players adopting S^m and S^s are closer to the trajectories when the players adopt S^* than when the set of linear strategies S^l is selected.

Figure 3.6 shows the regions in which each of the approximate strategies of player 1, namely u_1^l , u_1^m and u_1^s , yields the outcomes closest to the Nash equilibrium outcome, $J_1(u_1^*, u_2^*)$, when player 2 adheres to u_2^* , i.e it shows which of the approximate strategies minimises the difference $J_1(u_1^\bullet, u_2^*) - J_1(u_1^*, u_2^*)$ for the different initial conditions. Black, blue or red dots indicate that the strategies that achieve the minimum difference is u_1^l , u_1^m or u_1^s , respectively. For initial conditions where more than one strategy gives similar outcomes for player 1, in particular when the strategies result in the

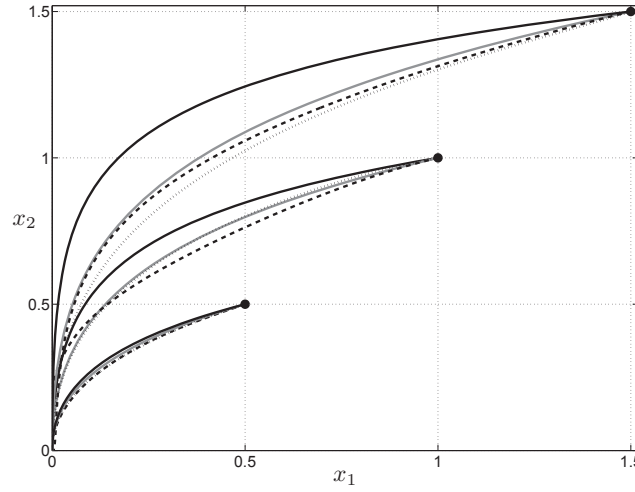


Figure 3.5: Trajectories of the state, x , when the players adopt the strategy pairs S^* (solid, gray line), S^l (solid, black line), S^m (dashed line) and S^s (dotted line) for three different x_0 (denoted by the solid circles).

outcomes differing by less than 0.0005, this is indicated by several dots around the corresponding initial conditions. Similarly, Figure 3.6 shows the regions in which each of the approximate strategies of player 2, i.e. u_2^l (black dots), u_2^m (blue dots) and u_2^s (red dots), yield the closest outcome to the Nash equilibrium outcome when player 1 adopts its Nash equilibrium strategy.

In all cases the state of the dynamic extensions ξ_1^m , ξ_2^m and ξ^s converge to zero, as expected. For completeness the time histories of the state of the dynamic extensions ξ_1^m (top) and ξ_2^m (middle) and ξ^s (bottom) for one of the the initial condition, namely $x_0 = (0, 2)^\top$, are shown in Figure 3.8.

The results indicate that when in the regions where the linear-quadratic approximation of the differential game is good, both dynamic strategies perform relatively well compared to the linear strategies. Outside this regions, the simulations indicate that any one of the two players loses less by deviating from its Nash equilibrium strategy to one of the dynamic strategies than what is the case when it adopts the linear strategy instead.

Note that these simulations are provided to give a qualitative illustration of the theory. Thus, it is possible that the performance of the dynamic feedback strategies can

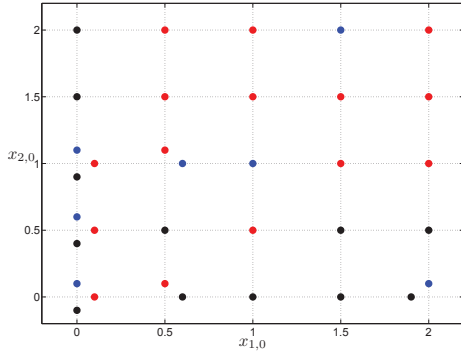


Figure 3.6: Regions where the approximate strategies u_1^l (black dots), u_1^m (blue dots) and u_1^s (red dots) result in outcomes closest to the Nash equilibrium outcome for player 1, when player 2 adheres to u_2^* .

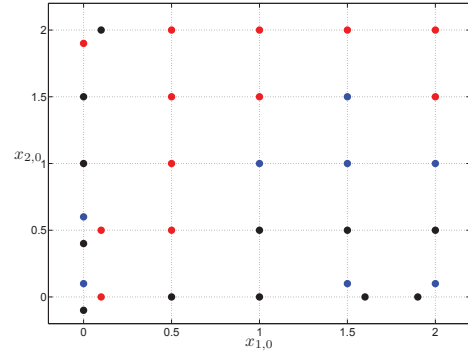


Figure 3.7: Regions where the approximate strategies u_2^l (black dots), u_2^m (blue dots) and u_2^s (red dots) result in outcomes closest to the Nash equilibrium outcome for player 1, when player 2 adheres to u_2^* .

be further improved by selecting the control parameters to optimise the performance for each initial condition.

3.7.2 Application to Lotka-Volterra Models

In this section a second numerical example illustrating the proposed methods is presented. A two-player differential game where the dynamical system satisfies the structural assumptions of Theorem 5 is studied. An *algebraic $\bar{P}$ matrix solution* is found and simulations with the dynamic controllers resulting from (3.34) with $i = 1, 2$, and (3.67) with $i = 1, 2$, are presented.

Consider the dynamical system

$$\begin{aligned}\dot{x}_1 &= b_1 x_1 - a_{11} x_1^2 - a_{21} x_1 x_2, \\ \dot{x}_2 &= b_2 x_2 - a_{12} x_1 x_2 - a_{22} x_2^2,\end{aligned}\tag{3.74}$$

where $x_1(t) \in \mathbb{R}$, $x_2(t) \in \mathbb{R}$, and b_i and a_{ij} , $i = 1, 2$, $j = 1, 2$ are positive parameters. These are the so-called *competitive Lotka-Volterra equations* and are commonly used to model systems consisting of competing species, including predator-prey systems [86]. The parameters b_1 and b_2 are the birth rates of the two species whereas a_{ij} , $i = 1, 2$ and $j = 1, 2$, are

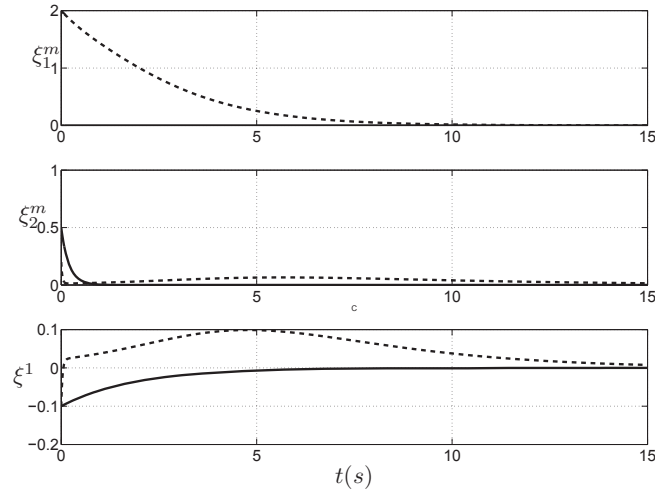


Figure 3.8: Time histories of states of the dynamic extensions ξ_1^m , ξ_2^m and ξ^s corresponding to the simulations for the initial condition $x_0 = (0, 2)^\top$. The solid and dashed lines denote the first and second components of the dynamic extensions, respectively.

predation efficiencies. The scalar parameters in (3.74) are typically assumed to be constants. We consider the somewhat different scenario in which the parameters a_{ii} and a_{ji} can be controlled around certain nominal values by player i , *i.e.* the predation efficiencies in (3.74) are given by $a_{ij} = a_{ij}^* + u_{ij}$, where a_{ij}^* are strictly positive constants, with $i = 1, 2$, $j = 1, 2$, and the vectors $u_1 = (u_{11}, u_{12})^\top$ and $u_2 = (u_{21}, u_{22})^\top$ are the control variables of players 1 and 2, respectively. In this case, the dynamical system (3.74) can be represented as an input-affine dynamical system (3.1) with $N = 2$, namely

$$\dot{x} = \begin{bmatrix} b_1 x_1 - a_{11}^* x_1^2 - a_{21}^* x_1 x_2 \\ b_2 x_2 - a_{21}^* x_1 x_2 - a_{22}^* x_2^2 \end{bmatrix} + \begin{bmatrix} -x_1^2 & 0 \\ 0 & -x_1 x_2 \end{bmatrix} u_1 + \begin{bmatrix} -x_1 x_2 & 0 \\ 0 & -x_2^2 \end{bmatrix} u_2, \quad (3.75)$$

where $f(x)$ has the equilibria

$$\begin{bmatrix} 0 & 0 \end{bmatrix}^\top, \begin{bmatrix} 0 & \frac{b_2}{a_{22}^*} \end{bmatrix}^\top, \begin{bmatrix} \frac{b_1}{a_{11}^*} & 0 \end{bmatrix}^\top, \begin{bmatrix} \frac{b_1 a_{22}^* - b_2 a_{21}^*}{-a_{12}^* a_{21}^* + a_{22}^* a_{11}^*} & \frac{b_2 a_{11}^* - b_1 a_{12}^*}{-a_{12}^* a_{21}^* + a_{22}^* a_{11}^*} \end{bmatrix}^\top.$$

Suppose that both species strive to drive the population to the latter of these equilibria, denoted by $x^* = [x_1^*, x_2^*]^\top$, while minimising its own control effort and maximising the efforts of the other species, *i.e.* each species attempts to minimise the cost functionals

(3.3), with running costs

$$q_1(x) = (x_2 - x_2^*)^2, \quad q_2(x) = (x_1 - x_1^*)^2. \quad (3.76)$$

Solving the differential game boils down to determining the *Nash equilibrium strategies* for the two players subject to the dynamics (3.75) and the cost functionals (3.3), with running costs (3.76).

Note that both $g_1(x)$ and $g_2(x)$ in (3.75) are full rank for all $x \neq 0$ and thus satisfy the structural assumptions of Theorem 5 and, as a result, dynamic approximations using either two *individual* dynamic extensions or a *shared* dynamic extension can be sought. Furthermore, the matrix-valued functions

$$P_1(\hat{x}) = \alpha \begin{bmatrix} 2 + 6\hat{x}_1 + 3\hat{x}_2 & 1 \\ 1 & 2 + 3\hat{x}_1 + 6\hat{x}_2 \end{bmatrix}, \quad (3.77)$$

$$P_2(\hat{x}) = \beta \begin{bmatrix} 2 + 6\hat{x}_1 + 3\hat{x}_2 & 1 \\ 1 & 2 + 3\hat{x}_1 + 6\hat{x}_2 \end{bmatrix}, \quad (3.78)$$

where $\hat{x} = [x_1 - x_1^*, x_2 - x_2^*]^\top$, constitute an *algebraic $\bar{P}$ matrix solution* to the problem for $\alpha > 0$ and $\beta > 0$ sufficiently large to ensure Σ_1 and Σ_2 in (3.22) are positive semidefinite.

Consider the case in which $b_1 = 1, b_2 = 1, a_{11}^* = 2, a_{22}^* = 2, a_{12}^* = 1$ and $a_{21}^* = 1$, yielding $x^* = \left[\frac{1}{3}, \frac{1}{3}\right]^\top$. Any selection of $\alpha \geq \frac{3}{10}$ and $\beta \geq \frac{3}{10}$ ensures the mappings (3.77) and (3.78) is an *algebraic $\bar{P}$ matrix solution* with $\Sigma_1 > 0$ and $\Sigma_2 > 0$. The system linearised about x^* is described by the matrices

$$A = \begin{bmatrix} -\frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} \end{bmatrix}, \quad B_1 = B_2 = \text{diag}\left(-\frac{1}{9}, -\frac{1}{9}\right), \quad (3.79)$$

and the solution to the corresponding coupled AREs (3.9) are

$$\bar{P}_1^l = \begin{bmatrix} 0.105 & -0.232 \\ -0.232 & 0.855 \end{bmatrix}, \quad \bar{P}_2^l = \begin{bmatrix} 0.855 & -0.232 \\ -0.232 & 0.105 \end{bmatrix}. \quad (3.80)$$

Note that $P_1 + P_2 > 0$. It follows that the Nash equilibrium strategies for the linearised system with the quadratic costs (3.76) are $u_i^l = -B_i^\top P_i^l x$, for $i = 1, 2$. The pair of strategies $\mathcal{S}^l = \{u_1^l, u_2^l\}$ is admissible and constitute the Nash equilibrium solution to the linear-quadratic approximation of the differential game.

In what follows dynamic approximations, using the two methods of Theorem 3 and Theorem 5 are found utilising (3.77) and (3.78), with $\alpha = \frac{3}{10}$ and $\beta = \frac{3}{10}$, as an *algebraic $\bar{P}$ matrix solution*. Finally, simulations for the two dynamic approximations, as well as the linear-quadratic approximation of the problem.

Let $\mathcal{S}^m = \{u_1^m, u_2^m\}$ denote⁹ the approximate dynamic solution which relies on the introduction of *individual* dynamic extensions, with dynamics given by (3.34). Consider next the case in which an approximate dynamic solution is sought using a *shared* dynamic extension, in accordance with Theorem 5. Note that the alternative conditions in Remark 10 are satisfied by Ψ_1 and Ψ_2 . Let $\mathcal{S}^s = \{u_1^s, u_2^s\}$ denote the approximate dynamic solution using a shared extension, where the dynamic control laws are given by (3.57).

Simulations have been run using the solutions to the linear-quadratic approximation, namely $\mathcal{S}^l$, and using the two approximate dynamic solutions $\mathcal{S}^m$ and $\mathcal{S}^s$, for different initial conditions $\hat{x}(0) = \hat{x}_0$. As in the previous example different parameters for the dynamic controllers have been used for the different initial conditions.

Figures 3.9 and 3.10 show the difference between the outcomes for player 1 and player, respectively, when the players adopt the set of dynamic strategies $\mathcal{S}^m$ and when the set of the linear strategies $\mathcal{S}^l$ is adopted, *i.e.* it shows $d_i^m = J_i(u_1^m, u_2^m) - J_i(u_1^l, u_2^l)$, $i = 1, 2$, for different initial conditions. Similarly, Figures 3.11 and 3.12 shows $d_i^s = J_i(u_1^s, u_2^s) - J_i(u_1^l, u_2^l)$ for $i = 1$ and $i = 2$, respectively. Let u_i^d , $i = 1, 2$, denote either one of the dynamic strategies and note that for all initial conditions both players obtain a more favourable outcome by adhering to the dynamic strategies, that is $J_i(u_1^d, u_2^d) - J_i(u_1^l, u_2^l) \leq 0$, $i = 1, 2$, holds for all $\hat{x}_0 = (\hat{x}_{0,1}, \hat{x}_{0,2})^\top$ considered in these simulations.

Figure 3.13 shows the trajectories of the state, x , when the players adopt the

⁹Again, for clarity of presentation the dynamic extensions are not included in the notation describing the dynamic feedback strategies.

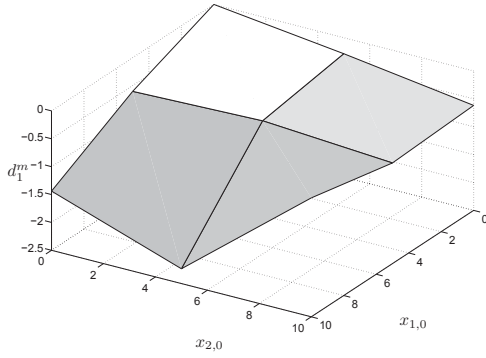


Figure 3.9: The values of d_1^m , corresponding to the different initial conditions, $\hat{x}_0$.

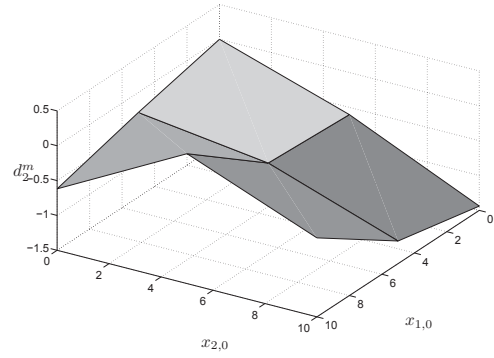


Figure 3.10: The values of $d_{2-0.5}^m$, corresponding to the different initial conditions, $\hat{x}_0$.

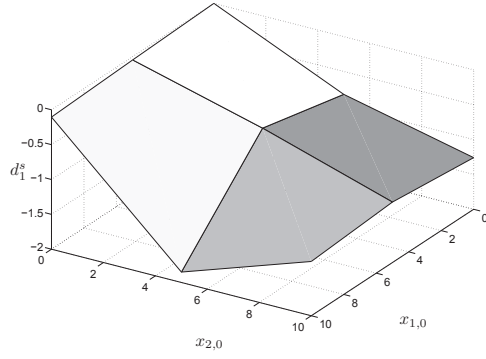


Figure 3.11: The values of d_1^s , corresponding to the different initial conditions, $\hat{x}_0$.

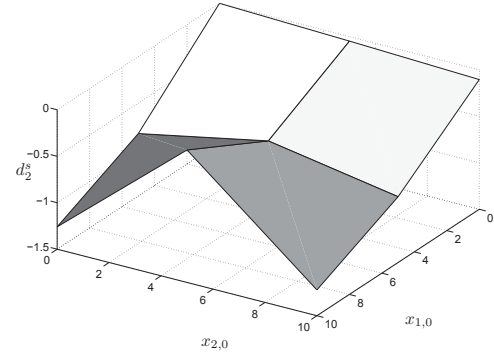


Figure 3.12: The values of d_2^s , corresponding to the different initial conditions, $\hat{x}_0$.

strategy pairs $\mathcal{S}^l$ (solid line), $\mathcal{S}^m$ (dashed line) and $\mathcal{S}^s$ (dotted line) for the initial condition $\hat{x}_0 = (5, 10)^\top$, which corresponds to the initial condition $x_0 = (5.33, 10.33)^\top$ in the original coordinates. The solid circle and square denote the initial and final values of the state x , respectively. For this particular initial condition, $R_1 = \text{diag}\{0.3, 0.2\}$, $R_2 = \text{diag}\{0.1, 0.1\}$, $k_1 = k_2 = 0.5$, $\xi_1^m(0) = [-0.1, -0.2]^\top$ and $\xi_2^m(0) = [0.08, 0.08]^\top$ have been selected for the dynamic strategies u_1^m and u_2^m . For the dynamic strategies u_1^s and u_2^s , on the other hand, $R_1 = \text{diag}(0.05, 0.1)$, $R_2 = \text{diag}(0.7, 0.8)$, $k = 0.5$ and $\xi_1^m(0) = [0.15, 0.1]^\top$ have been chosen. As can be seen, the state x converges to the equilibrium x^* in all three cases. Note that both the dynamic approximate solutions $\mathcal{S}^m$ and $\mathcal{S}^s$ result in very similar trajectories.

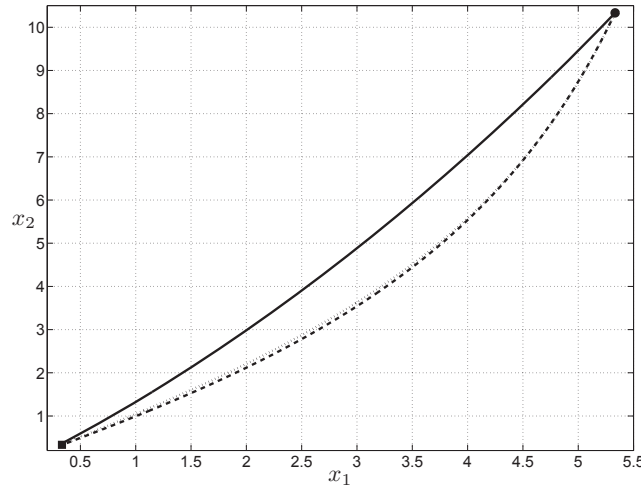


Figure 3.13: Trajectories of the state x when the players adopt the strategy pairs S^l (solid line), S^m (dashed line) and S^s (dotted line) for the initial condition $x_0 = (10.33, 5.33)^\top$ (denoted by the solid circle). The solid square denotes the final value of the state.

3.8 Applications to Multi-Agent Systems

In this Chapter a class of infinite-horizon nonzero-sum differential games has been discussed. For an N -player differential game within this class of problems, two ways of constructing dynamic strategies, using either N dynamic extensions or a shared dynamic extension have been presented. Furthermore, it has been shown that the dynamic strategies constitute local ϵ_α -Nash equilibrium strategies for the differential game and the two methods have been illustrated on two examples of differential games in Section 3.7.

As mentioned in Section 1.1 the agents in a multi-agent system can be thought of as players participating in a differential game, in which their goals may, or may not, be conflicting. In this section a differential game involving multi-agent systems is introduced and approximate solutions using the results of Section 3.6 are proposed. This work is motivated by three main factors. Firstly, along with the examples in Section 3.7 it serves to illustrate the theory discussed in Section 3.6. Furthermore, the cost functionals of the agents are not of the form (3.11) (or (3.3) in the 2-player case). Thus the problem illustrates the claims in Remark 1 and Proposition 1, *i.e.* it is demonstrated that the developed

theory can easily be adopted to problems with cost functionals different from (3.11) (or (3.3) in the 2-player case), provided certain conditions, stated in Proposition 1, are satisfied. Finally, the multi-agent collision avoidance problem highlights multi-agent systems as a possible application domain for the theory concerning differential games. This will be further developed in Chapter 5, in which the problem of continuously monitoring a region using a team of agents is discussed.

The problem of steering a team of agents from their initial positions to a predefined end-configuration while avoiding collisions is formulated as a differential game. This is referred to as the *multi-agent collision avoidance problem*. A method for obtaining approximate solutions for the differential game is then presented. This relies upon the conceptual results developed in Section 3.6, which provide ϵ_α -Nash strategies for general nonzero-sum differential games. Using this methods the *multi-agent collision avoidance problem* is solved. It is shown that approximate solutions guarantee that the task of reaching the final configuration while avoiding collisions is achieved. The theory is illustrated by simulations.

The remainder of the section is organised as follows. The problem is formulated in Section 3.8.1 and its solution is given in Sections 3.8.2 and 3.8.3. Simulations are presented in Section 3.8.4. Finally, some conclusions are drawn and directions for future work are identified in Section 7.2.

3.8.1 Problem Formulation

In this section we introduce and discuss the *multi-agent collision avoidance* problem, which consists in controlling a multi-agent system in such a way that each agent reaches a desired target position while simultaneously avoiding collisions between the other agents. Towards this end, consider a team of N agents moving in the plane, which are described by single-integrators, *i.e.*

$$\dot{x}_i = u_i, \quad (3.81)$$

for $i = 1, \dots, N$, with $x_i(t) \in \mathbb{R}^2$ and $u_i(t) \in \mathbb{R}^2$. The vectors x_i and u_i describe the position and the velocity of the i -th agent on the Cartesian plane, respectively.

Suppose that desired target positions x_i^* , $i = 1, \dots, N$, are assigned to each agent and let $\tilde{x}_i = x_i - x_i^*$ denote the error variable with respect to the desired configuration. In the following we assume that sufficient time is provided to accomplish the task of steering each agent from its initial condition, $x_i(0)$, to its target position, x_i^* . The above objective should be achieved while avoiding collisions between agents.

In particular, each agent is to maintain a certain distance between itself and the other members of the team. Thus, we define the *avoidance region* of the i -th agent as follows.

Definition 20. Given $r_i > 0$ and a time instant $\bar{t} \geq 0$, define the set $\mathcal{R}_i^{\bar{t}} = \{x : \|x_i(\bar{t}) - x\|^2 < r_i^2\}$, where $x \in \mathbb{R}^2$. The *avoidance region* of the i -th agent at $\bar{t}$, denoted $\Omega_i^{\bar{t}}$, is defined as $\Omega_i^{\bar{t}} = \bigcup_{j=1, \dots, N, j \neq i} \mathcal{R}_j$.

The parameter r_i plays the role of a *safety radius* for the i -th agent. Note that, since the team may consists of agents that differ in some way, *e.g.* they may have different sizes, different values of safety radius may be associated to each agent. Let $\bar{\Omega}_i^{\bar{t}}$ denote the complement of the set $\Omega_i^{\bar{t}}$. Then, a collision between the i -th and the j -th agent is defined as follows.

Definition 21. Given $r_i > 0$, the i -th agent is said to collide with the j -th agent if there exists a time instant $\bar{t} \geq 0$ such that $\|x_i(\bar{t}) - x_j(\bar{t})\|^2 < \max\{r_i^2, r_j^2\}$.

Finally, a collision-free trajectory for the agent i is defined as follows.

Definition 22. The i -th agent is said to be collision free if for all $\bar{t} \geq 0$, $x_i(\bar{t}) \notin \Omega_i^{\bar{t}}$, or equivalently for all $\bar{t} \geq 0$, $x_i(\bar{t}) \in \bar{\Omega}_i^{\bar{t}}$.

Remark 13. In the above definition a circular “geometry” is adopted, namely the set $\mathcal{R}_i^{\bar{t}}$ is a circular region centered at x_i . However, different geometries can easily be accounted for by modifying Definitions 20 and 21. In particular, the alternative definition $\mathcal{R}_i^{\bar{t}} = \{x : \|x_i(\bar{t}) - x\|_{M_i}^2 < r_i^2, \quad j = 1, \dots, N, \quad j \neq i\}$, where the positive-definite matrix $M_i \in \mathbb{R}^{2 \times 2}$ allows for “non-circular” geometries. ▲

Using the above discussion the multi-agent collision avoidance problem can be formulated as a non-cooperative differential game, as detailed in the following definition.

Problem 8. Consider a multi-agent system consisting of $N > 1$ agents with dynamics (3.81), for $i = 1, \dots, N$, and let $\tilde{x} = [\tilde{x}_1^\top, \dots, \tilde{x}_N^\top]^\top$. The *multi-agent collision avoidance problem* consists in steering each agent from its initial position to the predefined target position while avoiding inter-agent collisions and minimising its own effort. Thus, each agent attempts to minimise the cost-functional

$$J_i(u_1, \dots, u_N) = \frac{1}{2} \int_0^\infty (q_i(\tilde{x}) + \|u_i(t)\|^2) dt, \quad (3.82)$$

where $q_i(\tilde{x}) > 0$, with $q_i(0) = 0$, is a running cost given by

$$q_i(\tilde{x}) = (\alpha_i + \beta_i g_i(\tilde{x})) \tilde{x}_i^\top \tilde{x}_i, \quad (3.83)$$

with constants $\alpha_i > 0$, $\beta_i > 0$ and where $g_i(\tilde{x}) \geq 0$ is a barrier function penalising the i -th agent from approaching within a distance r_i of any other agent, namely such that

$$\lim_{x_i(t) \rightarrow \Omega_i^c} g_i(\tilde{x}) = +\infty.$$

Solving the multi-agent collision avoidance problem then boils down to determining the *Nash strategies* for each player, namely the set of strategies $\mathcal{S}^* = \{u_1^*, \dots, u_N^*\}$ satisfying

$$J_i(u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_i, \dots, u_N^*), \quad (3.84)$$

for all $u_i \neq u_i^*$, $i = 1, \dots, N$, and rendering the zero-equilibrium of the system locally asymptotically stable: each agent reaches its target position without entering its *avoidance region*.

In what follows an inverse barrier function, *i.e.*

$$g_i(\tilde{x}) = \sum_{j=1, j \neq i}^N \frac{1}{(\|(\tilde{x}_i + x_i^*) - (\tilde{x}_j + x_j^*)\|^2 - r_i^2)^c}, \quad (3.85)$$

where $c > 0$ is considered, although alternative definitions are possible.

In [87, 88] the problem of obtaining coverage of a region using a team of agents

while avoiding collisions has been solved by formulating the coverage problem as an optimal control problem and adding a posteriori a barrier function, penalising collisions, to the value functions corresponding to each agent. In contrast, we here include the collision avoidance in the formulation of the differential game.

Note that, as far as their primary objective of reaching x_i^* is concerned, the agents may be cooperating, *e.g.* for monitoring and surveillance, or competing.

3.8.2 Solving the Multi-Agent Collision Avoidance Problem

To solve Problem 8, the Hamilton-Jacobi-Isaacs equations associated with the differential game characterised by the cost functionals (3.82) and the dynamics (3.81), for $i = 1, \dots, N$, must be solved, namely

$$-\frac{1}{2} \frac{\partial V_i}{\partial x_i} \frac{\partial V_i}{\partial x_i}^\top + \frac{1}{2} q_i(x) - \sum_{j=1, j \neq i}^N \frac{\partial V_i}{\partial x_i} \frac{\partial V_j}{\partial x_j}^\top = 0, \quad (3.86)$$

with $V_i > 0$ and $V_i(0) = 0$, $i = 1, \dots, N$, [6, 10, 32]. Provided a solution to the system of PDEs can be found, the Nash equilibrium strategy of the i -th agent is given by

$$u_i^* = - \frac{\partial V_i(\tilde{x}_1, \dots, \tilde{x}_N)^\top}{\partial \tilde{x}_i}. \quad (3.87)$$

However, closed-form solutions to (3.86), $i = 1, \dots, N$, cannot be found and it is necessary to settle for an approximate solution of the differential game.

Earlier in this chapter two methods for solving approximately nonzero-sum differential games have been presented. In what follows the approximate dynamic solution relying on the introduction of a shared dynamic extension and the notion of *algebraic $\bar{P}$ matrix solution* is used to solve the multi-agent collision avoidance problem, *i.e.* Problem 8.

Since this problem deals with agents with single-integrator dynamics and cost functionals which are not of the form (3.11), an algebraic $\bar{P}$ matrix solution for Problem 8 is defined again for this particular differential game. This illustrates, as specified in

Remark 1, that the notion of Algebraic $\bar{P}$ matrix solution and the proposed methods of constructing dynamic strategies that constitute local ϵ_α -Nash equilibrium strategies for the differential game can be easily extended to differential games with cost functionals that are not of the form (3.11), or (3.3) in the 2-player case.

Using the machinery introduced in Section 3.4, and letting $\Sigma_i : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N \times 2N}$ such that $\Sigma_i(\tilde{x}) = \Sigma_i(\tilde{x})^\top \geq 0, i = 1, \dots, N$, for all $\tilde{x} \in \mathbb{R}^{2N}$, the matrix-valued functions $P_i(\tilde{x}) = P_i(\tilde{x})^\top, i = 1, \dots, N$, are said to constitute an *algebraic $\bar{P}$ matrix solution* for the differential game with Hamilton-Jacobi-Isaacs equations (3.86), for $i = 1, \dots, N$, if the following conditions hold.

(i) For all $x \in \mathbb{R}^n$ and $i = 1, \dots, N$,

$$-P_i(\tilde{x})P_i(\tilde{x}) + Q_i(\tilde{x}) - \sum_{j=1, j \neq i}^N 2P_i(\tilde{x})P_j(\tilde{x}) + \Sigma_i(\tilde{x}) = 0. \quad (3.88)$$

(ii) For $i = 1, \dots, N, P_i(0) = \bar{P}_i$, where $\bar{P}_i$ denotes the symmetric positive definite solution of the *coupled Riccati equations*

$$-\bar{P}_i\bar{P}_i - \sum_{j=1, j \neq i}^N (\bar{P}_i\bar{P}_j + \bar{P}_j\bar{P}_i) + \bar{Q}_i + \bar{\Sigma}_i = 0, \quad (3.89)$$

where $Q_i(\tilde{x})$ is such that $q_i(\tilde{x}) = \tilde{x}^\top Q_i(\tilde{x})\tilde{x}$ and $\bar{Q}_i = Q_i(0)$.

Using the *algebraic $\bar{P}$ matrix solution*, define the value functions as in (3.55), with $\xi \in \mathbb{R}^{2N}$, where $R_i = R_i^\top > 0, i = 1, \dots, N$. Furthermore, let $\xi = (\xi_1^\top, \dots, \xi_N^\top)^\top$, where $\xi_i \in \mathbb{R}^2, i = 1, \dots, N$. Recall that the partial derivatives of $V_i(\tilde{x}, \xi)$ are given by

$$\begin{aligned} \frac{\partial V_i(x)}{\partial x} &= \tilde{x}^\top P_i(\tilde{x}) + (\tilde{x} - \xi)^\top (R_i - \Phi_i(\tilde{x}, \xi))^\top, \\ \frac{\partial V_i(x)}{\partial \xi} &= \tilde{x}^\top \Psi_i(\tilde{x}, \xi)\tilde{x} - (\tilde{x} - \xi)R_i, \end{aligned} \quad (3.90)$$

where $\Phi_i(\tilde{x}, \xi)$ is such that $P_i(\tilde{x}) - P_i(\xi) = (\tilde{x} - \xi)\Phi_i(\tilde{x}, \xi)^\top$ and $\Psi_i(\tilde{x}, \xi)$ is the Jacobian matrix of the mapping $\frac{1}{2}P_i(\xi)\tilde{x}$ with respect to ξ . It follows from Theorem 6 (or Theorem 5 in the 2-player case), that there exists $k > 0$ and $R_i, i = 1, \dots, N$ such that the set of

dynamic strategies given by

$$\begin{aligned} u_i &= -P_i(\tilde{x})\tilde{x} + (R - \Phi_i(\tilde{x}, \xi))(\tilde{x} - \xi), \\ \dot{\xi} &= -k \sum_{i=1}^N \left(\Psi_i(\tilde{x}, \xi)^\top \tilde{x} - R_i(\tilde{x} - \xi) \right), \end{aligned} \quad (3.91)$$

constitute ϵ_α -Nash solutions for the differential game described by the cost-functionals (3.82) and the dynamics (3.81), $i = 1, \dots, N$. Thus, obtaining an approximate solution of the multi-agent collision avoidance problem that ensures all agents reach their targets without collisions occurring boils down to determining the *algebraic $\bar{P}$ matrix* solution, *i.e.* matrix-valued functions satisfying (3.88).

3.8.3 Multi-Agent Collision Avoidance

Approximate feedback solutions for the multi-agent collision avoidance problem are constructed here. It is assumed that the following conditions are satisfied by the initial and target configurations of the agents.

Assumption 5. The initial positions of the agents respect the safety radius of each agent, *i.e.*

$$\|x_i(0) - x_j(0)\| > r_i,$$

for all $i = 1, \dots, N, j = 1, \dots, N, j \neq i$.

Assumption 6. The target positions for each agent respect the safety radius of each agent, *i.e.*

$$\|x_i^* - x_j^*\| > r_i.$$

for all $i = 1, \dots, N, j = 1, \dots, N, j \neq i$,

In what follows the notation $A = [a_{ij}]$ is used as a shorthand for the matrix

$$A = \begin{bmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \vdots & a_{NN} \end{bmatrix}.$$

Consider the dynamic extension $\xi = (\xi_1, \dots, \xi_N) \in \mathbb{R}^{2N}$ and the matrix-valued functions $P_1(x) \in \mathbb{R}^{2N \times 2N}, \dots, P_N(x) \in \mathbb{R}^{2N \times 2N}$ given by

$$P_i(\tilde{x}) = [P_{ij}^i(x)]^\top + \gamma_i I, \quad (3.92)$$

where $P_{ij}^i \in \mathbb{R}^{2 \times 2}$, $i = 1, \dots, N$, $j = 1, \dots, N$ and $\gamma_i > 0$ is a constant parameter. Furthermore, $P_{ij}^i = 0$ for $j \neq i$ whereas

$$P_{ii}^i(\tilde{x}) = \left[\sqrt{\alpha_i + \beta_i q_i(\tilde{x})} I \right].$$

To clarify the notation, $P_1(\tilde{x})$ and $P_2(\tilde{x})$ are given in (3.93) and (3.94).

$$P_1 = \text{diag} \left(\sqrt{\alpha_1 + \beta_1 g_1(\tilde{x})} I, \gamma_1 I \right), \quad (3.93)$$

$$P_2 = \text{diag} \left(\gamma_2 I, (\sqrt{\alpha_2 + \beta_2 g_2(\tilde{x})} + \gamma_2) I, \gamma_2 I \right). \quad (3.94)$$

Finally, $R_i = [N_{ij}^i]$, where $N_{ij}^i \in \mathbb{R}^{2 \times 2}$, $i = 1, \dots, N$, $j = 1, \dots, N$. Using this notation the following theorem holds true.

Theorem 7. Consider Problem 8 and the algebraic $\bar{P}$ solution (3.92), $i = 1, \dots, N$, and suppose Assumptions 5 and 6 are satisfied. Then there exist R_i and $k_i > 0$, $i = 1, \dots, N$, such that the dynamic strategy

$$\begin{aligned} u_i &= -\tilde{x}_i \left(\sqrt{\alpha_i + \beta_i g_i(\xi)} + \gamma_i \right) - \sum_{p=1}^N N_{ip}^i (\tilde{x}_p - \xi_p) \\ \dot{\xi} &= -k \left(\frac{\beta_i \tilde{x}_i^\top \tilde{x}_i}{2\sqrt{\alpha_i + \beta_i g_i(\xi)}} \nabla_\xi g_i(\xi)^\top - R_i(\tilde{x} - \xi) \right), \end{aligned} \quad (3.95)$$

with $i = 1, \dots, N$, is such that $\lim_{t \rightarrow \infty} \tilde{x}_i(t) = 0$, $x_i(t) \in \bar{\Omega}_i^t$, for all $t \geq 0$ and $\lim_{t \rightarrow \infty} \xi(t) = 0$. Furthermore, the set of strategies (3.95), $i = 1, \dots, N$, constitutes an ϵ -Nash equilibrium for the differential game defined by the cost-functionals (3.82) and the system (3.81), $i = 1, \dots, N$. $\diamond$

Proof: The proof consists of two steps, the first one to show that the $P_i(\tilde{x})$, $i = 1, \dots, N$, give an *algebraic $\bar{P}$ matrix solution* for the differential game associated to Problem 8. It then follows from Theorem 6 (or Theorem 5 in the 2-player case), that the dynamic control strategies (3.95), $i = 1, \dots, N$, solve the problem approximately. Secondly, it is shown that provided Assumptions 5 and 6 are satisfied, the agents converge to the desired target points and collision avoidance is guaranteed along the trajectories of the agents.

First, note that it follows from (5.12) that $Q_i(\tilde{x}) = [a_{ij}]$, where $a_{ij} \in \mathbb{R}^2$, is such that $a_{ij} = 0$ for all $j \neq i$ and $a_{ii} = (\alpha_i + \beta_i g_i(\tilde{x}))I_2$. Then, clearly

$$\hat{P}_i(\tilde{x}) = \text{diag}\{0, \dots, 0, \sqrt{\alpha_i + \beta_i g_i(\tilde{x})}I, 0, \dots, 0\},$$

$i = 1, \dots, N$, satisfy (3.24) with $\Sigma_i(\tilde{x}) = 0$ for $i = 1, \dots, N$. Similarly, $P_1(\tilde{x}), \dots, P_N(\tilde{x})$, as defined in (3.92), satisfy (3.88) with

$$\Sigma_i(\tilde{x}) = \text{diag}\{\sigma_{11}, \sigma_{22}, \dots, \sigma_{NN}\} + \sum_{p=1}^N \gamma_i \gamma_p I,$$

where

$$\begin{aligned} \sigma_{ii} &= 2\gamma_i \sqrt{\alpha_i + \beta_i g_i(\tilde{x})} + \sum_{j=1, j \neq i}^N \gamma_j \sqrt{\alpha_i + \beta_i g_i(\tilde{x})}, \\ \sigma_{jj} &= \gamma_i \sqrt{\alpha_j + \beta_j g_j(\tilde{x})}. \end{aligned}$$

A direct substitution shows that the equations (3.89) are satisfied by (3.92), $i = 1, \dots, N$. It follows that the matrix-valued functions $P_1(\tilde{x}), \dots, P_N(\tilde{x})$ give an *algebraic $\bar{P}$ matrix solution* for the differential game associated to Problem 8. Furthermore, using (3.92), $i = 1, \dots, N$, as an algebraic solution, the dynamic strategies (3.67) are given by (3.95), $i = 1, \dots, N$.

It now remains to show that the resulting closed-loop trajectories are collision-free. The dynamic strategies (3.91) are the Nash-equilibrium strategies of a differential game with dynamics (3.81) and cost-functionals

$$J_i(u_1, \dots, u_N) = \frac{1}{2} \int_0^\infty (q_i(\tilde{x}) + \|u_i(t)\|^2 + c_i(\tilde{x}, \xi)) dt,$$

where $c_i(\tilde{x}, \xi) \geq 0, i = 1, \dots, N$. The value functions $V_1, \dots, V_N$ are such that

$$V_i(\tilde{x}(0)) = \frac{1}{2} \int_0^\infty (q_i(\tilde{x}) + \|u_i(t)\|^2 + c_i(\tilde{x}, \xi)) dt.$$

Provided Assumption 5 is satisfied, $q_i(\tilde{x})$ is bounded and it follows that $V_i(\tilde{x}(0))$ is bounded as well. Since an infinite-horizon problem is considered, it is necessary that $q_i(0) = 0, i = 1, \dots, N$, for the problem to be well-posed. This can only be the case provided Assumption 6 is satisfied.

Finally, taking $W(\tilde{x}) = V_1(\tilde{x}) + \dots + V_N(\tilde{x})$ as a candidate Lyapunov function yields $\dot{W} \leq -\frac{1}{2} \sum_{i=1}^N q_i(\tilde{x})$. It follows that provided Assumption 5 is satisfied, hence $W(\tilde{x}(0))$ is bounded, and Assumption 6 is also satisfied, hence the problem is well-posed, the zero equilibrium of the closed-loop system is locally asymptotically stable: the agents converge to their target position. Furthermore, since $\dot{W} \leq 0$ and Assumption 5 holds, it follows that the agents reach their targets without entering their avoidance region. $\square$

Remark 14. The initial condition of the dynamic extension variable, $\xi(0)$, is of importance for the solution of the differential game: it must be selected so that $g_i(\xi(0))$ is bounded for all $i = 1, \dots, N$. $\blacktriangle$

Remark 15. The algebraic $\bar{P}$ matrix solutions (3.92), $i = 1, \dots, N$ can be modified such that $\Sigma_i(\tilde{x})$ contains the barrier function $g_i(\tilde{x})$. This may be useful to satisfy technical assumptions discussed in Remark 6. $\blacktriangle$

3.8.4 Simulations

In this section simulations to illustrate the theory are presented. While the function (3.85) *ideally* behaves as a barrier function, in a discrete-step implementation - intrinsically required by numerical integration - the agents may miss, and consequently cross, the barrier due to the time discretization. Therefore, to render the implementation robust, the following alternative definition of the barrier function is used

$$g_i(\tilde{x}) = \frac{1}{(\max\{0, \|(\tilde{x}_i + x_i^*) - (\tilde{x}_j + x_j^*)\|^2 - r_i^2\})^c}. \quad (3.96)$$

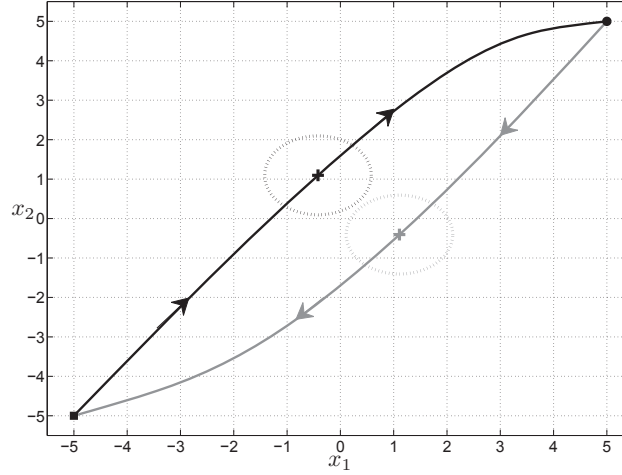


Figure 3.14: Trajectories of the first (dark line) and second (gray line) agent with $x_1(0) = [-5, -5]^\top$, $x_2(0) = [5, 5]^\top$, $x_1^* = [5, 5]^\top$ and $x_2^* = [-5, -5]^\top$. The square and circular markers indicate the initial conditions of $x_1(0)$ and $x_2(0)$, respectively.

Simulations have been run for different scenarios. First, two cases with $N = 2$ are considered. Then a case in which there are three agents that need to be coordinated on the plane is considered. Identical agents, with $r_i = 1 = r_s$, $\gamma_i = 1$, $\alpha_i = 1$, $\beta_i = 0.1$, $i = 1, \dots, N$, are considered in all simulations. For the barrier functions (3.96), the value $c = 1$ is used.

Consider first the case in which $N = 2$, $x_1(0) = [-5, -5]^\top$, $x_2(0) = [5, 5]^\top$, $x_1^* = x_2(0)$ and $x_2^* = x_1(0)$. Both agents apply the dynamic strategies (3.95), $i = 1, 2$, with $k = 1$, $R_1 = R_2 = I$ and $\xi(0) = [8, -3, -3, 8]^\top$. Figure 3.14 shows the trajectories of the first (dark line) and second (gray line) agents. The arrows indicate the direction of travel. The square and circular markers denote the initial positions of the first and second agent, respectively. The dotted lines denote the circumferences of the regions $\mathcal{R}_i$ of each agent at the time instant $t_c \geq 0$ in which the agents are closest to each other. The crosses indicate where this occurs on the trajectories of the agents. Note that these regions are not intersecting. Figure 3.15 shows the time history of $\|x_1(t) - x_2(t)\|$ (dark line) together with the value r_s (dashed line). As can be seen from Figure 3.14 both agents reach their targets and Figure 3.15 shows that the avoidance regions of the two agents never overlap while doing so. The middle and bottom graphs of Figure 3.15 display the time histories

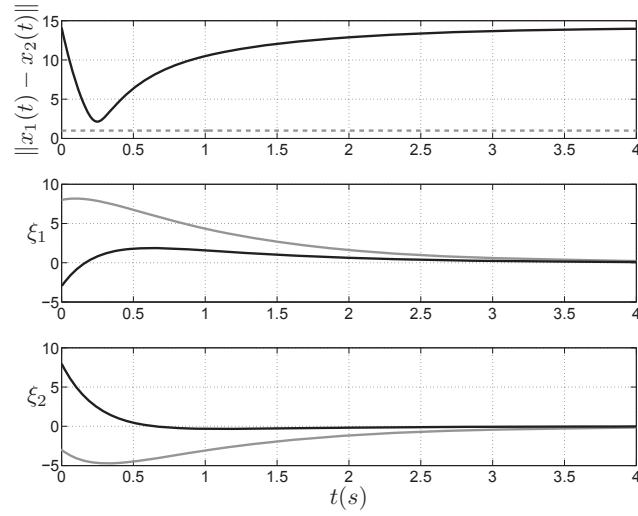


Figure 3.15: Top graph: time history of the Euclidean distance between the two agents, *i.e.* $\|x_1(t) - x_2(t)\|$, (solid line), together with r_s (dashed line). Middle graph: time histories of the components of the dynamic extension $\xi_1(t)$, with $\xi_1(0) = [8, -3]^\top$. Bottom graph: time histories of the components of the dynamic extension $\xi_2(t)$, with $\xi_2(0) = [-3, 8]^\top$.

of the components of the dynamic extensions $\xi_1(t)$ and $\xi_2(t)$, respectively. Figure 3.16 shows the time histories of the control strategies u_1 (top) and u_2 (bottom) adopted by agents 1 and 2, respectively. In both cases the solid line denotes the component of the strategies along the x -axis, whereas the dashed line denotes the components along the y -axis.

Consider a similar situation with two agents, this time with initial positions $x_1(0) = [-5, -5]^\top$ and $x_2(0) = [5, 5]^\top$ with target positions $x_1^* = [3, 2]^\top$ and $x_2^* = [-1, 1]^\top$. In this case $k = 1$, $R_1 = R_2 = I$ and $\xi(0) = [-20, 20, 20, -20]^\top$. Figure 3.17 shows the trajectories of the agents (black and gray lines, respectively). In this figure circular markers are used to denote the initial positions of the agents whereas the squares denote their target position, x_i^* , $i = 1, 2$. As in the previous example both agents reach the desired final destinations without colliding. The top graph of Figure 3.18 shows the time history of $\|x_1(t) - x_2(t)\|$, whereas the middle and bottom graphs show the time histories of the components of $\xi_1(t)$ and $\xi_2(t)$, respectively. Figure 3.19 displays the time histories of the components of the strategies u_1 (top) and u_2 (bottom) of agents 1 and 2, respectively.

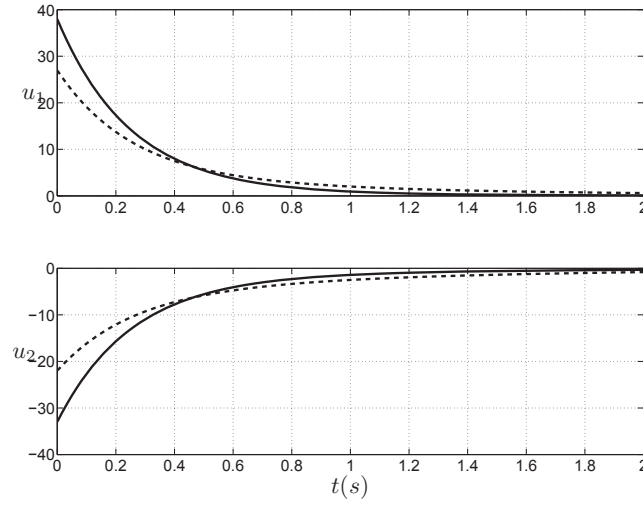


Figure 3.16: Time histories of the control strategies adopted by agents 1 (top) and 2 (bottom).

Solid and dashed lines indicate the components of the strategies along the x -axis and y -axis, respectively.

Finally, consider the scenario in which three agents are to be coordinated in the plane. More specifically, the primary deployment task consists in steering the agents from the vertexes of the triangle plotted with the dotted line in Figure 3.20 towards the vertexes of the, smaller, triangle plotted with the dash-dotted line. The initial conditions are $x_1(0) = [-2, 0]^\top$, $x_2(0) = [-1, 4]^\top$ and $x_3(0) = [3, 3]^\top$, whereas the targets are $x_1^* = [-1, 3]^\top$, $x_2^* = [1, 1.5]^\top$ and $x_3^* = [-1, 0]^\top$. In this simulation $k = 12$, $R_1 = R_2 = I$ and $\xi(0) = [40, -40, -40, 40, 40, 40]^\top$. Figure 3.20 shows the trajectories of the first (dark line), second (dark-gray line) and third (light-gray line) agent, together with the triangles describing the initial (dotted line) and final configurations (dash-dotted line). Figure 3.21 displays the time histories of the distances between each pair of agents, namely $d_{ij}(t) \triangleq \|x_i(t) - x_j(t)\|$. Figure 3.19 shows the time histories of the control strategies u_1 (top), u_2 (middle) and u_3 (bottom) adopted by agents 1, 2 and 3, respectively. Solid and dashed lines indicate the components of the strategies along the x -axis and y -axis, respectively.

The three scenarios considered here show that the agents reach their target positions while avoiding collisions.

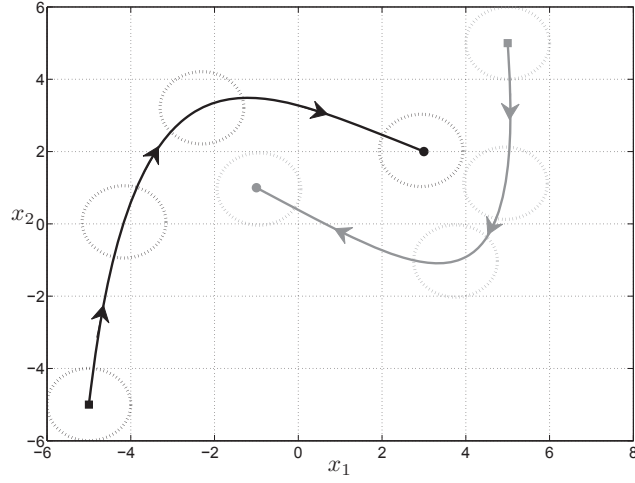


Figure 3.17: Trajectories of the first (dark line) and second (gray line) agent with $x_1(0) = [-5, -5]^\top$ and $x_2(0) = [5, 5]^\top$. The circular and square markers indicate the initial and target positions, $x_1^* = [3, 2]^\top$ and $x_2^* = [-1, 1]^\top$, of the agents, respectively.

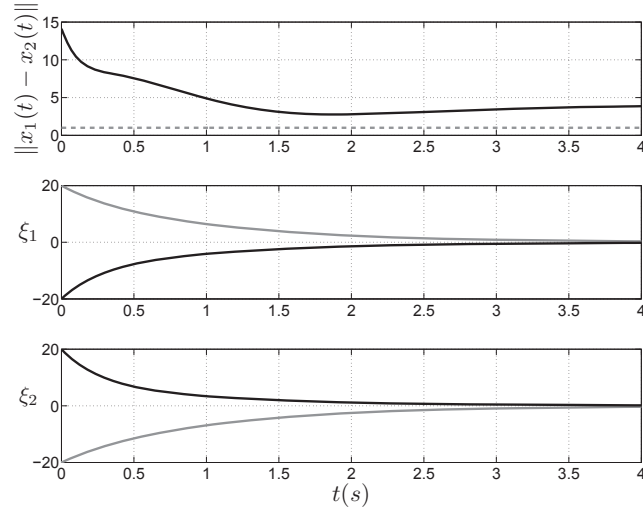


Figure 3.18: Top graph: time history of the Euclidean distance between the two agents (solid line) together with r_s (dashed line). Middle graph: time histories of the components of the dynamic extension $\xi_1(t)$, with $\xi_1(0) = [20, -20]^\top$. Bottom graph: time histories of the components of the dynamic extension $\xi_2(t)$, with $\xi_2(0) = [-20, 20]^\top$.

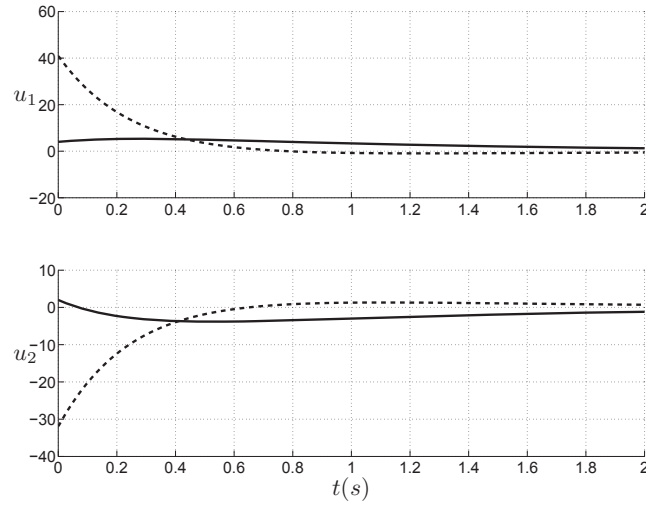


Figure 3.19: Time histories of the control strategies adopted by agents 1 (top) and 2 (bottom).

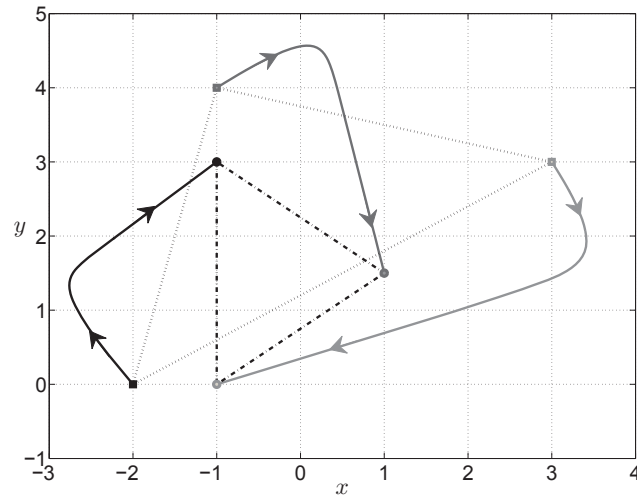


Figure 3.20: Trajectories of the first (dark line), second (dark gray line) and third (light gray line) agent, together with the triangles describing the initial (dotted line) and final, desired, configurations (dot-dashed line).

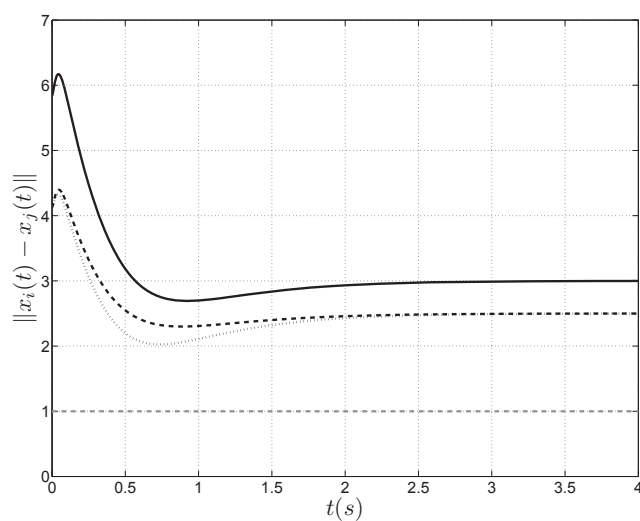


Figure 3.21: Time histories of the Euclidean distances $d_{12}(t)$ (solid line), $d_{13}(t)$ (dotted line) and $d_{23}(t)$ (dashed line) with $d_{ij}(t) \triangleq \|x_i(t) - x_j(t)\|$.

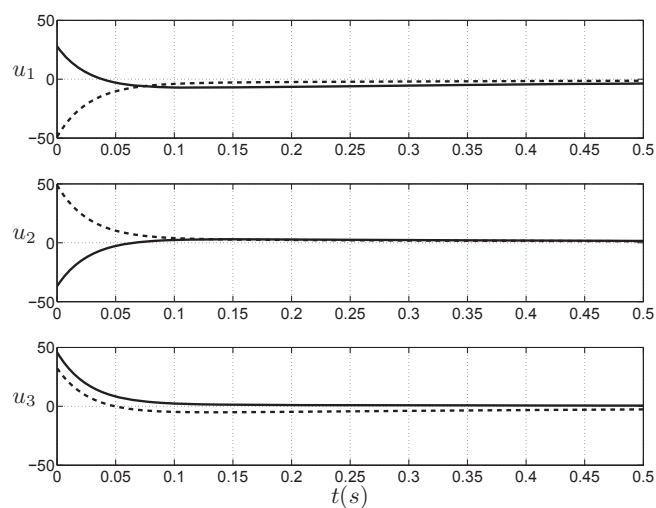


Figure 3.22: Time histories of the control strategies adopted by agents 1 (top) and 2 (middle) and 3 (bottom).

3.9 Conclusion

In this chapter classes of nonlinear, nonzero-sum differential games with input-affine dynamics are studied, with focus on feedback Nash equilibria. The notions of α -admissible strategies and ϵ_α -Nash equilibrium strategies are introduced and it is shown that for linear-quadratic differential games solving matrix inequalities in place of the algebraic Riccati equations yields ϵ_α -Nash equilibrium solutions. This observation is then extended to the nonlinear setting. Without solving the Hamilton-Jacobi-Isaacs PDEs associated with the differential games directly, dynamic strategies that instead satisfy partial differential *inequalities* are identified, yielding local ϵ_α -Nash equilibrium solutions. The proposed method relies on dynamic state feedback and the notion of *algebraic $\bar{P}$ matrix solution*. Either *several* or a single, *shared* dynamic extension can be used, depending on the structural properties of the dynamic system.

The proposed methods are applied to two illustrative examples of nonzero-sum differential games. The first example is one for which the Nash equilibrium solution is known and it is shown that both approximate dynamic solutions result in an outcome closer to the Nash equilibrium outcome for *both* the players than the linear-quadratic approximation of the problem does. Then a *Lotka-Volterra* model with two competing species is presented. Simulations indicate that both approximate dynamic solutions, using *several* and a *shared* dynamic extensions, result in an improvement compared to the performance resulting from the linear-quadratic approximation of the problem. Finally, the problem of steering a team of N agents from their initial position to a desired end-configuration while avoiding collisions, *i.e.* the *multi-agent collision avoidance problem*, is considered. The problem is formulated as a differential game which is solved using the approximate methods presented earlier in the chapter: an *algebraic $\bar{P}$ matrix solution* for the *multi-agent collision avoidance problem* is identified and used to design dynamic control laws that guarantee that each agent reaches its final destination while avoiding collisions with other agents, provided the initial and final configurations satisfy certain conditions. The theory is supported by simulations.

Chapter 4

2-Player Stackelberg Differential Games

4.1 Introduction

One of the most common solution concepts associated with game theoretic problems is that of *Nash equilibrium solutions*, discussed in Chapters 2 and 3, whereby it is assumed that all players are rational and announce their strategies simultaneously [6, 10, 70]. A different solution concept was introduced by Stackelberg in 1934 [89]. In Stackelberg differential games some decision makers are able to act prior to the other players which then react in a rational manner, *i.e.* the players act in a specific order [6]. Thus, unlike Nash equilibria, Stackelberg equilibria introduce a *hierarchy* between the players. Some simple examples of the different types of games are the game of “rock-paper-scissor”, where all players act simultaneously (Nash), and “tic-tac-toe” or chess, where the players act in a specific order (Stackelberg). In some cases the order in which the players act is irrelevant, in which case the Nash and Stackelberg equilibria are equivalent. This is, however, not generally the case. As mentioned in Chapter 2 the concepts introduced when focusing on feedback Nash equilibrium strategies can be extended to different solution concepts related to differential games as well. In this chapter a simple 2-player differential game is considered for which *Stackelberg equilibrium strategies* are defined and

Stackelberg equilibrium solutions are sought. Different solution concepts for Stackelberg differential games exist, namely *open-loop*, *closed-loop* and *feedback* Stackelberg solutions. These solution concepts are discussed in [90]. Unlike closed-loop Stackelberg solutions, feedback Stackelberg solutions can be found via dynamic programming, and these are studied in this chapter.

In [89] feedback Stackelberg solutions of an infinite-horizon stochastic differential game are considered. In this chapter a similar 2-player deterministic differential game is introduced. As when determining Nash equilibrium solutions, obtaining feedback Stackelberg equilibrium solutions for a given differential game relies on the solution of a system of coupled PDEs. The Stackelberg equilibrium solution to a linear-quadratic differential game, similar to the linear-quadratic differential game introduced in Section 2.3.1, relies on solutions of coupled algebraic Riccati equations. In Chapter 3 two methods of obtaining strategies that constitute local ϵ_α -Nash equilibrium strategies for a class of nonzero-sum differential games have been introduced. In what follows, the method developed in Section 3.6, namely the approach which relies on a dynamic extension which is shared among the players, is extended to the case in which feedback Stackelberg solutions are sought. In particular a two-player nonzero-sum differential game, with cost functionals similar to those considered in [89] is considered. The results can easily be extended to problems with more general cost functionals, similar to those considered in Chapter 3, and to problems with $N > 2$ players.

The motivation behind this chapter is twofold. Firstly, it demonstrates that the methods introduced in Chapter 3 are applicable to solution concepts other than Nash equilibrium solutions. Secondly, it is the first step towards considering differential games with more complicated information patterns. For example, such a *non-standard information pattern* would occur in the situation in which a group of players has information about the strategies of a second group of players whereas the second group do not have access to similar information regarding the actions of the first group of players.

The remainder of the chapter is organised as follows. In Section 4.2 a formal definition of the problem and its solution is given. An *algebraic $\bar{P}$ matrix Stackelberg solution*,

which is similar in spirit to the algebraic $\bar{P}$ matrix solution defined in Chapter 3, is defined in Section 4.3. A method for constructing approximate solutions is presented in Section 4.4, in which also the notion of the approximation is made precise. Simulations illustrating the method are then given in Section 4.5. Finally, some concluding remarks are given in Section 4.6.

4.2 The Two-Player Stackelberg Differential Game

Consider the input-affine dynamical system

$$\dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2, \quad (4.1)$$

where $x(t) \in \mathbb{R}^n$, with $n > 0$, is the state of the system, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $g_1(x) \in \mathbb{R}^{n \times m}$ and $g_2(x) \in \mathbb{R}^{n \times m}$, with $m > 0$, are smooth mappings and $u_1(t) \in \mathbb{R}^m$ and $u_2(t) \in \mathbb{R}^m$ are the control signals of player 1 and 2, respectively.

Assumption 7. The origin of $\mathbb{R}^n$ is an equilibrium point of the vector field $f(x)$, i.e. $f(0) = 0$.

Recall, as remarked in Section 3.2, that a consequence of Assumption 7 is that we can write $f(x) = F(x)x$, for some (not unique) matrix-valued function $F(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$.

Suppose the game is such that the players announce their strategies in a pre-defined order. In particular suppose that player 1 acts before player 2: in this case player 1 is referred to as the *leader*, whereas player 2 is referred to as the *follower*. Similarly to [91], the leader seeks to minimise a cost functional of the form

$$J_1(x(0), u_1, u_2) \triangleq \frac{1}{2} \int_0^\infty \left(q_1(x(t)) + 2\theta u_1(t)^\top u_2(t) + \|u_1(t)\|^2 \right) dt, \quad (4.2)$$

whereas the follower seeks to minimise a cost functional of the form

$$J_2(x(0), u_1, u_2) \triangleq \frac{1}{2} \int_0^\infty \left(q_2(x(t)) + 2\theta u_2(t)^\top u_1(t) + \|u_2(t)\|^2 \right) dt, \quad (4.3)$$

where $\theta \in \mathbb{R}$, and the running costs are such that $q_1(x) = x^\top Q_1(x)x \geq 0$, $q_2(x) = x^\top Q_2(x)x \geq 0$, where $Q_1 \in \mathbb{R}^{n \times n}$ and $Q_2 \in \mathbb{R}^{n \times n}$, and $q_1(x) + q_2(x) > 0$ for all $x \neq 0$. The cost functionals (4.2) and (4.3) are such that the agents seek to minimise their own running costs and control efforts. In addition, if $\theta > 0$, the players seek to minimise the cross-product between their control strategies (or maximise the product if instead $\theta < 0$).

The cost functionals remain bounded if $J_1 + \lambda J_2 \geq 0$ for some $\lambda \in \mathbb{R}^+$. This is guaranteed provided θ and λ are such that $\lambda - \theta^2(1 + \lambda)^2 \geq 0$. The range of allowable θ is maximised by the selection $\lambda = 1$ and the resulting bound is given in the following statement, which is assumed to hold in the remainder of the paper.

Assumption 8. The parameter θ , weighting the cross-product between the two control inputs, is such that $|\theta| \leq \frac{1}{2}$.

Remark 16. The range of θ in Assumption 8 is more restrictive than that in [91]. We consider this range of θ to ensure boundedness of $J_1 + J_2$. However, in [91] a discount factor is included in the running cost, which eliminates the “problem” of boundedness, hence allowing for a larger range of θ . ▲

Remark 17. Cost functionals different from (4.2) and (4.3) can be considered, provided they are such that there exists $\lambda > 0$ such that $J_1 + \lambda J_2$ is bounded. The cost functionals (4.2) and (4.3) have been selected to allow for a straight-forward comparison with the results in [91]. ▲

In what follows *feedback Stackelberg equilibria* for the differential game are considered. Admissible strategies are defined similarly to what has been done in Definition 7 as follows.

Definition 23. A pair of state feedback strategies $\mathcal{S} = \{u_1, u_2(u_1)\}$ is said to be admissible for the non-cooperative Stackelberg differential game if it renders the zero equilibrium of the closed-loop system (locally) asymptotically stable.

Problem 9. Consider the system (4.1) and the cost functionals of the leader and the follower, namely (4.2) and (4.3), respectively. The 2-player Stackelberg differential game

consists in determining a pair of admissible feedback strategies, $\mathcal{S}^* = \{u_1^*, u_2^*(u_1^*)\}$ satisfying

$$\begin{aligned} J_1(x(0), u_1^*, u_2^*(u_1^*)) &\leq J_1(x(0), u_1, u_2^*(u_1)), \\ J_2(x(0), u_1^*, u_2^*(u_1^*)) &\leq J_2(x(0), u_1^*, u_2(u_1^*)), \end{aligned} \quad (4.4)$$

for all admissible $\mathcal{S}_1 = \{u_1, u_2^*(u_1)\}$ and $\mathcal{S}_2 = \{u_1^*, u_2(u_1^*)\}$, with $u_1 \neq u_1^*$ and $u_2(u_1^*) \neq u_2^*(u_1^*)$.

The pair of strategies $\mathcal{S}^*$ is said to be a Stackelberg equilibrium solution of the two-player Stackelberg differential game.

Remark 18. Since there is an ordering between the two players, the Stackelberg strategy of the follower, *i.e.* $u_2(u_1^*)$, depends on the strategy of the leader, *i.e.* u_1^* . This is not the case when solving for Nash equilibria, where both players act simultaneously. $\blacktriangle$

The Hamiltonians associated with players 1 and 2 are

$$\begin{aligned} H_1(x, u_1, u_2, \lambda_1) &= \frac{1}{2} \left(q_1(x) + 2\theta u_1^\top u_2 + \|u_1\|^2 \right) + \lambda_1^\top (f(x) + g_1(x)u_1 + g_2(x)u_2), \\ H_2(x, u_1, u_2, \lambda_1) &= \frac{1}{2} \left(q_2(x) + 2\theta u_2^\top u_1 + \|u_2\|^2 \right) + \lambda_2^\top (f(x) + g_1(x)u_1 + g_2(x)u_2), \end{aligned}$$

where λ_1 and λ_2 are the co-states. Feedback Stackelberg equilibria are such that the Hamiltonians are minimised [91]. Knowing the action of the leader, the follower's response is

$$u_2^*(u_1) = \arg \min_{u_2} H_2(x, u_1, u_2, \lambda_2) = -g_2^\top \lambda_2 - \theta u_1. \quad (4.5)$$

Anticipating this behaviour, the leader should select its strategy according to

$$u_1^* = \arg \min_{u_1} H_1(x, u_1, u_2^*, \lambda_1) = -\frac{1}{1 - 2\theta^2} \left((g_1(x) - \theta g_2(x))^\top \lambda_1 - \theta g_2(x)^\top \lambda_2 \right). \quad (4.6)$$

Let $\Delta_g(x) = g_1(x) - \theta g_2(x)$ and consider the PDEs

$$\begin{aligned} \frac{\partial V_1}{\partial x} f(x) + \frac{1}{2} q_1(x) - \frac{\partial V_1}{\partial x} g_2(x) g_2(x)^\top \frac{\partial V_2}{\partial x}^\top - \frac{1}{2} \frac{1}{1-2\theta^2} \left\| \frac{\partial V_1}{\partial x} \Delta_g(x) - \theta \frac{\partial V_2}{\partial x} g_2(x) \right\|^2 = 0, \\ \frac{\partial V_2}{\partial x} f(x) + \frac{1}{2} q_2(x) + \frac{1}{1-2\theta^2} \theta \frac{\partial V_2}{\partial x} g_2(x) g_2(x)^\top \frac{\partial V_2}{\partial x}^\top - \frac{1}{1-2\theta^2} \frac{\partial V_2}{\partial x} g_1(x) \Delta_g(x)^\top \frac{\partial V_1}{\partial x}^\top \\ - \frac{1}{2} \frac{1}{(1-2\theta^2)^2} \left\| (1-\theta^2) \frac{\partial V_2}{\partial x} g_2(x) - \theta \frac{\partial V_1}{\partial x} \Delta_g(x) \right\|^2 = 0, \end{aligned} \quad (4.7)$$

with solutions V_i such that $V_i(0) = 0$, for $i = 1, 2$, for all $x \neq 0$. Then, the Stackelberg equilibrium solution for the game defined by the dynamics (4.1) and the cost functionals (4.2) and (4.3) with player 1 as leader and player 2 as follower, *i.e.* Problem 9, is $\mathcal{S}^* = \{u_1^*, u_2^*(u_1^*)\}$, where the strategies of each player are given by

$$\begin{aligned} u_1^* &= \frac{-1}{1-2\theta^2} \left(\Delta_g(x)^\top \frac{\partial V_1}{\partial x}^\top - \theta g_2(x) \frac{\partial V_2}{\partial x}^\top \right), \\ u_2^*(u_1^*) &= -g_2(x)^\top \frac{\partial V_2}{\partial x}^\top - \theta u_1^*, \end{aligned} \quad (4.8)$$

provided the set of strategies $\mathcal{S}^*$ is admissible.

Note that admissibility of $\mathcal{S}^*$ follows if $W = V_1(x) + V_2(x) > 0$, for all $x \neq 0$. Moreover, $\dot{W} < 0$, for all $x \neq 0$, is then implied by the PDEs (4.7). Thus, it follows from standard Lyapunov arguments that the feedback strategies (4.8) are admissible if the solution of (4.7) is such that $W > 0$.

Consider now the *linear-quadratic approximation of the problem*. Recall that in a neighbourhood of the origin the system (4.1) can be approximated by a linear system (3.6) with matrices A , B_1 and B_2 given by (3.7), and the running costs $q_i(x)$ can be approximated as quadratic costs, namely $x^\top \bar{Q}_i x$, where $\bar{Q}_i = Q_i(0)$, for $i = 1, 2$. As seen in Section 3.2.1 for the resulting linear-quadratic differential game, when focusing on linear feedback strategies only, the PDEs (4.7), reduce to coupled AREs. For the Stackelberg

differential game considered here the coupled AREs are

$$\begin{aligned}
& P_1 A + A^\top P_1 + \bar{Q}_1 - P_1 B_2 B_2^\top P_2 - P_2 B_2 B_2^\top P_1 \\
& - \frac{1}{1 - 2\theta^2} (P_1 \Delta_B - \theta P_2 B_2) (\Delta_B^\top P_1 - \theta B_2^\top P_2) = 0, \\
& P_2 A + A^\top P_2 + Q_2 + \frac{1}{1 - 2\theta^2} (2\theta P_2 B_2 B_2^\top P_2 - P_1 \Delta_B B_1^\top P_2 - P_2 B_1 \Delta_B^\top P_1) \\
& - \frac{1}{(1 - 2\theta^2)^2} ((1 - \theta^2) P_2 B_2 - \theta P_1 \Delta_B) ((1 - \theta^2) B_2^\top P_2 - \theta \Delta_B^\top P_1) = 0,
\end{aligned} \tag{4.9}$$

where $\Delta_B = (B_1 - \theta B_2)$. Suppose $P_1 = P_1^\top$ and $P_2 = P_2^\top$, such that $P_1 + P_2 > 0$, solve (4.9). Then, the Stackelberg equilibrium solution for the linear-quadratic differential game is $\mathcal{S}^* = \{u_1^*, u_2^*(u_1^*)\}$, where

$$\begin{aligned}
u_1^* &= \frac{-1}{1 - 2\theta^2} (\Delta_B^\top \bar{P}_1 - \theta B_2^\top \bar{P}_2) x, \\
u_2^*(u_1^*) &= -B_2^\top \bar{P}_2 x - \theta u_1^*.
\end{aligned}$$

Note that the by assumption $W = \frac{1}{2} x^\top (P_1 + P_2) x > 0$, for all $x \neq 0$. Furthermore, $\dot{W} < 0$, for all $x \neq 0$, are implied by (4.9). It follows that the pair of strategies $\mathcal{S}^*$ is admissible.

Remark 19. In [91] a stochastic linear-quadratic differential game with $B_1 = B_2 = I$ has been considered. The above results are consistent with [91]. $\blacktriangle$

Remark 20. When $\theta = 0$, the feedback Stackelberg equilibrium strategies and the feedback Nash equilibrium strategies coincide, *i.e.* in this case the order in which the players act is irrelevant to the corresponding outcomes [91, 92]. $\blacktriangle$

To conclude this section define the notions of α -admissible strategies and ϵ_α -Stackelberg equilibrium solutions, similarly to what has been done in Chapter 3.

Definition 24. The pair of strategies $\mathcal{S} = \{u_1, u_2(u_1)\}$ is said to be α -admissible for the non-cooperative Stackelberg differential game if the zero equilibrium of the system (4.1) in closed-loop with $\mathcal{S}$ is (locally) asymptotically stable and $\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of the closed-loop system around the origin.

Definition 25. Admissible strategies u_1^* and $u_2^*(u_1^*)$ are said to be ϵ_α -Stackelberg equilibrium strategies for player 1 and 2, respectively, of the Stackelberg differential game with dynamics (4.1) and cost functionals (4.2) and (4.3), with player 1 as leader and player 2 as follower, if for the pair of strategies $\mathcal{S}^* = \{u_1^*, u_2^*(u_1^*)\}$ there exists a non-negative constant $\epsilon_{x_0, \alpha}$, parametrised with respect to $x(0) = x_0$, and $\alpha > 0$ such that

$$J_1(x_0, u_1^*, u_2^*(u_1^*)) \leq J_1(x_0, u_1, u_2^*(u_1)) + \epsilon_{x_0, \alpha},$$

$$J_2(x_0, u_1^*, u_2^*(u_1^*)) \leq J_2(x_0, u_1^*, u_2(u_1^*)) + \epsilon_{x_0, \alpha},$$

for all α -admissible strategy pairs $\mathcal{S}_1 = \{u_1, u_2^*(u_1)\}$ and $\mathcal{S}_2 = \{u_1^*, u_2(u_1^*)\}$, with $u_1 \neq u_1^*$ and $u_2(u_1^*) \neq u_2^*(u_1^*)$.

A method for obtaining approximate solutions to the Stackelberg differential game is presented in what follows. The approach is similar in spirit to the approximate dynamic solution using a shared dynamic extension and the notion of *algebraic $\bar{P}$ matrix solutions* introduced in Section 3.6.2 for Nash differential games

4.3 Algebraic $\bar{P}$ Matrix Stackelberg Solution

The method for obtaining approximate solutions to the Stackelberg differential game in Problem 9 relies on the notion of *algebraic $\bar{P}$ matrix Stackelberg solution*, which is defined, similarly to the algebraic $\bar{P}$ matrix solution introduced in Definition 15, in this section.

Definition 26. Consider the system (4.1), the cost functionals (4.2) and (4.3) and the resulting Stackelberg differential game, with player 1 as leader and player 2 as follower. Let $\Sigma_1(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ and $\Sigma_2(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ be matrix-valued functions satisfying $\Sigma_i(x) > 0$, for all $x \in \mathbb{R}^n \setminus \{0\}$, and $\Sigma_i(0) = \bar{\Sigma}_i$, for $i = 1, 2$. The $\mathcal{C}^1$ matrix-valued functions $P_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $P_i(x) = P_i(x)^\top$, $i = 1, 2$, are said to be an $\mathcal{X}$ -algebraic $\bar{P}$ matrix solution¹ of the equations (4.7), if the following conditions are satisfied.

¹Provided the set $\mathcal{X}$ contains the origin.

i) For all $x \in \mathcal{X}$

$$\begin{aligned} & P_1(x)F(x) + F(x)^\top P_1(x) + Q_1(x) - \frac{1}{1-2\theta^2} \left\| P_1(x)\Delta_g(x) - \theta P_2(x)g_2(x) \right\|^2 \\ & - 2P_1(x)g_2(x)g_2(x)^\top P_2(x) + \Sigma_1(x) = 0, \end{aligned} \quad (4.10)$$

and

$$\begin{aligned} & P_2(x)F(x) + F(x)^\top P_2(x) + Q_2(x) - \frac{1}{(1-2\theta^2)^2} \left\| (1-\theta^2)P_2(x)g_2(x) - \theta P_1(x)\Delta_g(x) \right\|^2 \\ & + \frac{2}{1+2\theta^2} \left(\theta P_2(x)g_2(x)g_2(x)^\top P_2(x) \right) - \frac{1}{1+2\theta^2} \left(P_2(x)g_1(x)\Delta_g(x)^\top P_1(x) \right) \\ & - \frac{1}{1+2\theta^2} \left(P_1(x)\Delta_g(x)g_1(x)^\top P_2(x) \right) + \Sigma_2(x) = 0. \end{aligned} \quad (4.11)$$

ii) $P_1(0) = \bar{P}_1$ and $P_2(0) = \bar{P}_2$, where $\bar{P}_1$ and $\bar{P}_2$ are symmetric matrices, such that $\bar{P}_1 + \bar{P}_2 > 0$, and satisfy the coupled AREs

$$\begin{aligned} & \bar{P}_1 A + A^\top \bar{P}_1 + \bar{Q}_1 - \bar{P}_1 B_2 B_2^\top - \bar{P}_2 B_2 B_2^\top \bar{P}_2 \\ & - \frac{1}{1-2\theta^2} (\bar{P}_1 \Delta_B - \theta \bar{P}_2 B_2) (\Delta_B^\top \bar{P}_1 - \theta B_2^\top \bar{P}_2) + \bar{\Sigma}_1 = 0, \\ & \bar{P}_2 A + A^\top \bar{P}_2 + \bar{Q}_2 + \frac{1}{1-2\theta^2} \left(2\theta \bar{P}_2 B_2 B_2^\top \bar{P}_2 - \bar{P}_1 \Delta_B B_1^\top \bar{P}_2 - \bar{P}_2 B_1 \Delta_B^\top \bar{P}_1 \right) \\ & - \frac{1}{(1-2\theta^2)^2} \left((1-\theta^2) \bar{P}_2 B_2 - \theta \bar{P}_1 \Delta_B \right) ((1-\theta^2) B_2^\top - \theta \Delta_B^\top \bar{P}_1) + \bar{\Sigma}_2 = 0. \end{aligned} \quad (4.12)$$

If $\mathcal{X} = \mathbb{R}^n$, $P_1(x)$ and $P_2(x)$ are said to be an *algebraic $\bar{P}$ matrix Stackelberg solution*.

In what follows it assumed that *algebraic $\bar{P}$ matrix Stackelberg solutions* exist.

4.4 Approximate Solutions Using a Shared Dynamic Extension

Consider the Stackelberg differential game in Problem 9. Similar to what has been done in Section 3.6.2 a modified problem which approximates the differential game is formulated by introducing a dynamic extension, $\xi(t)$, which is shared by the two players. The notion of algebraic $\bar{P}$ matrix Stackelberg solutions is used to define, in the extended state-space,

functions which satisfy Hamilton-Jacobi-Isaacs partial differential inequalities. These functions are used to design dynamic feedback strategies that constitute (local) ϵ_α -Stackelberg equilibrium strategies.

In what follows we consider dynamic feedback strategies of the form

$$u_1 = \beta_1(x, \xi), \quad u_2 = \beta_2(x, \xi, \beta_1(x, \xi)), \quad \dot{\xi} = \tau(x, \xi), \quad (4.13)$$

with $\xi(t) \in \mathbb{R}^\nu$, for some $\nu > 0$, where τ , β_1 and β_2 are smooth mappings with $\tau(0, 0) = 0$, $\beta_1(0, 0) = 0$ and $\beta_2(0, 0, 0) = 0$.

Definition 27. The pair of dynamic feedback strategies $\mathcal{S} = \{u_1, u_2(u_1), \dot{\xi}\}$ is said to be admissible (α -admissible) if the zero equilibrium of the closed-loop system (4.1)-(4.13) is (locally) asymptotically stable ($\sigma(A_{cl} + \alpha I) \subset \mathbb{C}^-$, where A_{cl} is the matrix describing the linearisation of (3.1)-(3.26) around the origin).

Problem 10. Consider the system (4.1) and the cost functionals of the leader and the follower, *i.e.* (4.2) and (4.3). The *approximate dynamic Stackelberg differential game* consists in determining a pair of dynamic feedback strategies $\mathcal{S} = \{u_1, u_2(u_1), \dot{\xi}\}$ of the form (4.13), with $\xi(t) \in \mathbb{R}^n$, and non-negative functions $c_1 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $c_2 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, such that, for any $x(0)$, $\xi(0)$ and for any admissible $\mathcal{S}_1 = \{u_1, \beta_2(u_1), \dot{\xi}\}$, with $u_1 \neq \beta_1$, and $\mathcal{S}_2 = \{\beta_1, u_2(\beta_1), \dot{\xi}\}$, with $u_2 \neq \beta_2$,

$$\tilde{J}_1(x(0), \beta_1, \beta_2(\beta_1)) \leq \tilde{J}_1(x(0), u_1, \beta_2(u_1)),$$

$$\tilde{J}_2(x(0), \beta_1, \beta_2(\beta_1)) \leq \tilde{J}_2(x(0), \beta_1, u_2(\beta_1)),$$

where the modified cost functionals $\hat{J}_1$ and $\hat{J}_2$ are given by

$$\tilde{J}_i \triangleq \frac{1}{2} \int_0^\infty \left(q_i(x(t)) + 2\theta u_1(t)^\top u_2(t) + \|u_i(t)\|^2 + c_i(x(t), \xi(t)) \right) dt, \quad (4.14)$$

for $i = 1, 2$.

Suppose $P_1(x)$ and $P_2(x)$ constitute an *algebraic $\bar{P}$ matrix Stackelberg solution* and define

the functions

$$\begin{aligned} V_1(x, \xi) &= \frac{1}{2}x^\top P_1(\xi)x + \frac{1}{2}\|x - \xi\|_{R_1}^2, \\ V_2(x, \xi) &= \frac{1}{2}x^\top P_2(\xi)x + \frac{1}{2}\|x - \xi\|_{R_2}^2, \end{aligned} \quad (4.15)$$

where $R_1 \in \mathbb{R}^{n \times n}$ and $R_2 \in \mathbb{R}^{n \times n}$ are symmetric and positive-definite matrices, i.e. $R_1 = R_1^\top > 0$ and $R_2 = R_2^\top > 0$.

Notation similar to that introduced in Chapter 3 is used here, and is stated at this stage for clarity. Let $\Phi_1(x) \in \mathbb{R}^{n \times n}$ be such that $x^\top (P_1(x) - P_1(\xi)) = (x - \xi)^\top \Phi_1(x, \xi)^\top$ and similarly let $\Phi_2(x) \in \mathbb{R}^{n \times n}$ be such that $x^\top (P_2(x) - P_2(\xi)) = (x - \xi)^\top \Phi_2(x, \xi)^\top$. Furthermore let $\Psi_1(x, \xi)$ and $\Psi_2(x, \xi)$ denote the Jacobian matrices, with respect to ξ , of $\frac{1}{2}P_1(\xi)x$ and $\frac{1}{2}P_2(\xi)x$, respectively. The next statement provides a method for constructing a solution for Problem 10.

Theorem 8. Consider the system (4.1) and the cost functionals (4.2) and (4.3) with $\theta \in (-\frac{1}{2}, \frac{1}{2})$. Let $P_1(x)$ and $P_2(x)$ be an algebraic $\bar{P}$ matrix Stackelberg solution, satisfying $\bar{\Sigma}_1 > 0$ and $\bar{\Sigma}_2 > 0$, and let R_1 and R_2 be such that

$$R_i(R_1 + R_2) + (R_1 + R_2)R_i \geq 0, \quad (4.16)$$

for $i = 1, 2$. Then there exist $\bar{k} \geq 0$ and a neighbourhood of the origin $\Omega \subset \mathbb{R}^n \times \mathbb{R}^n$ such that, for all $k \geq \bar{k}$, the functions (4.15) solve the system of partial differential inequalities

$$\begin{aligned} \mathcal{HJ}_1 &= \frac{\partial V_1}{\partial x} f(x) + \frac{1}{2}q_1(x) + \frac{\partial V_1}{\partial \xi} \dot{\xi} - \frac{\partial V_1}{\partial x} g_2(x)g_2(x)^\top \frac{\partial V_2}{\partial x}^\top \\ &\quad - \frac{1}{2} \frac{1}{1 - 2\theta^2} \left\| \frac{\partial V_1}{\partial x} \Delta_g(x) - \theta \frac{\partial V_2}{\partial x} g_2(x) \right\|^2 \leq 0, \end{aligned} \quad (4.17)$$

and

$$\begin{aligned} \mathcal{HJ}_2 &= \frac{\partial V_2}{\partial x} f(x) + \frac{1}{2}q_2(x) + \frac{1}{1 - 2\theta^2} \theta \frac{\partial V_2}{\partial x} g_2(x)g_2(x)^\top \frac{\partial V_2}{\partial x}^\top - \frac{1}{1 - 2\theta^2} \frac{\partial V_2}{\partial x} g_1(x)\Delta_g(x)^\top \frac{\partial V_1}{\partial x}^\top \\ &\quad + \frac{\partial V_2}{\partial \xi} \dot{\xi} - \frac{1}{2} \frac{1}{(1 - 2\theta^2)^2} \left\| (1 - \theta^2) \frac{\partial V_2}{\partial x} g_2(x) - \theta \frac{\partial V_1}{\partial x} \Delta_g(x) \right\|^2 \leq 0, \end{aligned} \quad (4.18)$$

with $\dot{\xi} = -k \left(\frac{\partial V_1}{\partial \xi} + \frac{\partial V_2}{\partial \xi} \right)^\top$ and for all $(x, \xi) \in \Omega$. Furthermore, the dynamical system

$$\begin{aligned} u_1 = & -\frac{\Delta_g(x)^\top}{1-2\theta^2} \left(P_1(x)x + (R_1 - \Phi_1(x, \xi))(x - \xi) \right)^\top \\ & + \frac{\theta g_2(x)^\top}{1-2\theta^2} \left(P_2(x)x + (R_2 - \Phi_2(x, \xi))(x - \xi) \right)^\top, \\ u_2(u_1) = & -g_2(x)^\top \left(P_2(x)x + (R_2 - \Phi_2(x, \xi))(x - \xi) \right)^\top - \theta u_1, \end{aligned} \quad (4.19)$$

$$\dot{\xi} = -k \left(\frac{\partial V_1}{\partial \xi} + \frac{\partial V_2}{\partial \xi} \right)^\top,$$

is such that the set of dynamic feedback strategies $\mathcal{S} = \{u_1, u_2(u_1), \dot{\xi}\}$ is admissible and solves Problem 10, with $c_1(x, \xi) = -2\mathcal{H}\mathcal{J}_1(x, \xi)$ and $c_2(x, \xi) = -2\mathcal{H}\mathcal{J}_2(x, \xi)$. Finally, there exists a neighbourhood of the origin in which the set of dynamic feedback strategies $\mathcal{S} = \{u_1, u_2(u_1), \dot{\xi}\}$ constitutes an ϵ_α -Stackelberg equilibrium solution of Problem 9 for all $\alpha > 0$. $\diamond$

Proof: The proof consists of three parts. First it is shown that the functions (4.15) satisfy the partial differential inequalities (4.17) and (4.18). It is then demonstrated that the zero-equilibrium of the closed-loop system (4.1)-(4.19) is (locally) asymptotically stable and solves Problem 10. The final part of the claim is then proved exploiting Definitions 24 and 25.

The partial derivatives of the functions defined in (4.15) with respect to x and ξ are

$$\begin{aligned} \frac{\partial V_1}{\partial x} &= x^\top P_1(x) + (x - \xi)^\top (R_1 - \Phi_1(x, \xi)^\top), \\ \frac{\partial V_1}{\partial \xi} &= x^\top \Psi_1(x, \xi) - (x - \xi)^\top R_1, \\ \frac{\partial V_2}{\partial x} &= x^\top P_2(x) + (x - \xi)^\top (R_2 - \Phi_2(x, \xi)^\top), \\ \frac{\partial V_2}{\partial \xi} &= x^\top \Psi_2(x, \xi) - (x - \xi)^\top R_2. \end{aligned}$$

Let $\Upsilon_1 = \Delta_g^\top (R_1 - \Phi_1) - \theta g_2 (R_2 - \Phi_2)$. Using the properties of algebraic $\bar{P}$ Stackelberg

solutions, the inequality (4.17) can be written as a *quadratic* form

$$-\begin{bmatrix} x^\top & (x - \xi)^\top \end{bmatrix} (M_1 + kD_1) \begin{bmatrix} x \\ (x - \xi) \end{bmatrix} \leq 0. \quad (4.20)$$

The matrix M_1 is given by

$$M_1 = \begin{bmatrix} \frac{1}{2}\Sigma_1 & \Gamma_1^1 \\ \Gamma_1^1 & \Gamma_2^1 \end{bmatrix},$$

where

$$\Gamma_1^1 = \frac{1}{2} \left((P_2 g_2 g_2^\top - F^\top)(R_1 - \Phi_1) + P_1 g_2 g_2^\top (R_2 - \Phi_2) + \frac{1}{1 - 2\theta^2} (P_1 \Delta_g - \theta P_2 g_2) \Upsilon_1 \right),$$

and $\Gamma_2^1 = (R_1 - \Phi_1 g_2 g_2^\top (R_2 - \Phi_2)^\top + \frac{1}{2} \frac{1}{1 - 2\theta^2} \Upsilon_1^\top \Upsilon_1$, whereas the matrix D_1 is given by

$$D_1(x, \xi) = \begin{bmatrix} \Lambda_{11}^1 & \Lambda_{12}^1 \\ \Lambda_{12}^{1^\top} & \Lambda_{22}^1 \end{bmatrix},$$

where $\Lambda_{11}^1 = \Psi_1(\Psi_1 + \Psi_2)^\top + (\Psi_1 + \Psi_2)\Psi_1^\top$, $\Lambda_{12}^1 = -\Psi_1(R_1 + R_2) - (\Psi_1 + \Psi_2)R_1$ and $\Lambda_{22}^1 = R_1(R_1 + R_2) + (R_1 + R_2)R_1$. Since $\Psi_1(0, \xi) = 0$ and $\Psi_2(0, \xi) = 0$ it follows that $D_1|_{(0, \xi)} = \text{diag}\{0, \Lambda_{22}^1\}$, namely the inequalities (4.16), $i = 1, 2$, are such that $D_1|_{(0, \xi)}$ is positive semidefinite. Furthermore, $Z_1 = [I, 0]^\top$ spans the kernel of $D_1|_{(0, \xi)}$ and $Z_1^\top M_1|_{(0, 0)} Z_1 = \bar{\Sigma}_1 > 0$ and by continuity this product is positive definite in a neighbourhood of the origin. By [85] it follows that there exists $\bar{k}_1 > 0$ and a non-empty set Ω_1 containing the origin such that the inequality (4.17) is satisfied for all $k > \bar{k}_1$ and $(x, \xi) \in \Omega_1$.

Similary, the second inequality (4.18) can be written as

$$-\begin{bmatrix} x^\top & (x - \xi)^\top \end{bmatrix} (M_2 + kD_2) \begin{bmatrix} x \\ (x - \xi) \end{bmatrix} \leq 0, \quad (4.21)$$

where

$$M_2 = \begin{bmatrix} \frac{1}{2}\Sigma_2 & \Gamma_1^2 \\ \Gamma_1^2 & \Gamma_2^2 \end{bmatrix},$$

with

$$\begin{aligned}\Gamma_1^2 = & \frac{1}{2} \left(\frac{1}{1-2\theta^2} (P_1 \Delta_g g_1^\top - 2\theta P_2 g_2 g_2^\top) - F^\top \right) (R_2 - \Phi_2) \\ & + \frac{1}{2} \left(\frac{1}{1-2\theta^2} P_2 g_1 \Delta_g^\top (R_1 - \Phi_1) + \Upsilon_2^\top \Upsilon_2 \right),\end{aligned}$$

and

$$\Gamma_2^2 = \frac{1}{2} \Upsilon_2^\top \Upsilon_2 + (R_2 - \Phi_1^\top) \left(\frac{1}{1-2\theta^2} g_1 \Delta_g^\top (R_1 - \Phi_1) - g_2 g_2^\top (R_1 - \Phi_2) \right),$$

where $\Upsilon_2 = \frac{1}{1-2\theta^2} ((1-\theta^2)g_2(R_2 - \Phi_2) - \theta \Delta_g^\top (R_1 - \Phi_1^\top))$, and

$$D_2(x, \xi) = \begin{bmatrix} \Lambda_{11}^2 & \Lambda_{12}^2 \\ \Lambda_{12}^{2^\top} & \Lambda_{22}^2 \end{bmatrix},$$

with $\Lambda_{11}^2 = \Psi_2(\Psi_1 + \Psi_2)^\top + (\Psi_1 + \Psi_2)\Psi_2^\top$, $\Lambda_{12}^2 = -\Psi_2(R_1 + R_2) - (\Psi_1 + \Psi_2)R_2$ and $\Lambda_{22}^2 = R_2(R_1 + R_2) + (R_1 + R_2)R_2$. As with $D_1(x, \xi)$, $D_2|_{(0, \xi)} = \text{diag}\{0, \Lambda_{22}^1\}$, $Z_2 = [I, 0]^\top$ spans the kernel of $D_2|_{(0, \xi)}$ and $Z_2^\top M_2|_{(0,0)} Z_2 = \bar{\Sigma}_2 > 0$. By continuity $Z_2^\top M_2|_{(0,0)} Z_2 > 0$ in a neighbourhood of the origin, which implies that there exists $\bar{k}_2 > 0$ and a set Ω_2 containing the origin such that for all $k \geq \bar{k}_2$ and all $(x, \xi) \in \Omega_2$ the inequality (4.18) is satisfied.

Let $\bar{k} = \max\{\bar{k}_1, \bar{k}_2\}$ and $\Omega = \Omega_1 \cap \Omega_2$. The above implies that the inequalities (4.17) and (4.18) are satisfied for all $k \geq \bar{k}$ and all $(x, \xi) \in \Omega$, proving the first part of the statement.

Let $W = V_1 + \lambda V_2$ for some $\lambda > 0$ and note that (at least in a neighbourhood of the origin) $W > 0$, for all $(x, \xi) \neq 0$, by definition of *algebraic $\bar{P}$ matrix Stackelberg solutions*. It follows from the inequalities (4.17) and (4.18) that, for all $(x, \xi) \in \Omega$,

$$\begin{aligned}\dot{W} \leq & -\frac{1}{2} (q_1(x) + \lambda q_2(x)) \\ & - \frac{1}{2} \begin{bmatrix} \Xi_1^\top & \Xi_2^\top \end{bmatrix} \begin{bmatrix} 1 & -(1+\lambda)\theta \\ -(1+\lambda)\theta & \lambda \end{bmatrix} \begin{bmatrix} \Xi_1 \\ \Xi_2 \end{bmatrix},\end{aligned}\tag{4.22}$$

where

$$\Xi_1 = \frac{1}{1 - 2\theta^2} \left(\Delta_g^\top \frac{\partial V_1}{\partial x} - \theta g_2^\top \frac{\partial V_2}{\partial x} \right)$$

and

$$\Xi_2 = \frac{\theta}{1 - 2\theta^2} \Delta_g^\top \frac{\partial V_1}{\partial x} - (1 + \theta) g_2^\top \frac{\partial V_2}{\partial x}.$$

Thus, for all $(x, \xi) \in \Omega \setminus \{0\}$, $\dot{W} \leq 0$ by Assumption 8. By standard Lyapunov arguments and by Lasalle's invariance principle, it follows that $\lim_{t \rightarrow \infty} x(t) = 0$. Furthermore, local asymptotic stability of the zero equilibrium of the system $\dot{\xi} = -k \left(\frac{\partial V_1}{\partial x}^\top + \frac{\partial V_2}{\partial x}^\top \right) \Big|_{x=0}$ follows from (4.16). Consequently, local asymptotic stability of the zero equilibrium, *i.e.* $(x, \xi) = (0, 0)$ follows by standard arguments on interconnected systems. As a result the dynamic feedback strategies solve Problem 10 with $c_1(x, \xi) = -2\mathcal{H}\mathcal{J}_1$ and $c_2(x, \xi) = -2\mathcal{H}\mathcal{J}_2$. Finally, similarly to what has been shown in Chapter 3, for any α -admissible control strategy the integrals $\int_0^\infty c_i(x(t), \xi(t)) dt$, $i = 1, 2$ are bounded in a neighbourhood $\bar{\Omega}$, which may depend on α , of the origin. It follows that $\mathcal{S}$ is an ϵ_α -Stackelberg equilibrium solution for Problem 9. $\square$

4.5 Simulations

In this section a numerical example illustrating the theory is presented. Consider the dynamical system

$$\dot{x} = \begin{bmatrix} (1 + x_2^2) & 2 \frac{x_1 x_2}{1 + x_2^2} \\ 0 & 1 \end{bmatrix} (u_1 + u_2), \quad (4.23)$$

and suppose player 1 is the leader and player 2 is the follower, seeking to minimise the cost functionals (4.2) and (4.3), respectively, where the running costs are given by

$$q_1(x) = a_1 \frac{x_1^2}{1 + x_2^2} + a_2 x_2^2,$$

$$q_2(x) = b_1 \frac{x_1^2}{1 + x_2^2} + b_2 x_2^2,$$

with $a_i \geq 0$ and $b_i \geq 0$, for $i = 1, 2$. Suppose $a_1 = 18$, $a_2 = 8$, $b_1 = 9$, $b_2 = 4$ and $\theta = 0.5$.

The matrix-valued functions

$$P_1(x) = \text{diag} \left(\frac{\alpha_1}{(1+x_2^2)^2}, \alpha_2(1+c_1x_1^2+c_2x_2^2) \right),$$

and

$$P_2(x) = \text{diag} \left(\frac{\beta_1}{(1+x_2^2)^2}, \beta_2(1+c_3x_1^2+c_4x_2^2) \right),$$

with $c_i \geq 0, i = 1, \dots, 4$, $\alpha_1\beta_1 \geq \max\left\{\frac{1}{2}a_1, \beta_1^2 + \frac{1}{2}b_1\right\}$ and $\alpha_2\beta_2 \geq \max\left\{\frac{1}{2}a_2, \beta_2^2 + \frac{1}{2}b_2\right\}$, $\beta_1 \leq \alpha_1$ and $\beta_2 \leq \alpha_2$, constitute an *algebraic $\bar{P}$ matrix Stackelberg solution*, with $\Sigma_1(x) > 0$ and $\Sigma_2(x) > 0$, for the differential game. Let $\alpha_1 = 4.5$, $\alpha_2 = 3$, $\beta_1 = 2.5$, $\beta_2 = 2$, $c_i = 0$, $i = 1, \dots, 4$ and let u_1^d and $u_2^d(u_1^d)$ denote² the corresponding dynamic feedback strategies given by (4.19).

By performing the change of coordinates $\hat{x}_1 = x_1(1+x_2)^2$ and $\hat{x}_2 = x_2$, the problem can be transformed into a linear-quadratic Stackelberg differential game, with $A = 0$, $B_1 = B_2 = I$, $\bar{Q}_1 = \text{diag}(18, 8)$ and $\bar{Q}_2 = \text{diag}(9, 4)$, for which $\hat{P}_1 = \text{diag}(4, 2)$ and $\hat{P}_2 = \text{diag}(2, 2)$ solve the coupled AREs (4.9). It follows that the Stackelberg equilibrium strategies (in the original coordinates) are

$$\begin{aligned} u_1^* &= - \left[\frac{6x_1}{1+x_2^2}, 2x_2 \right]^\top, \\ u_2^*(u_1^*) &= - \left[2\frac{x_1}{1+x_2^2}, 2x_2 \right]^\top - \theta u_1^*, \end{aligned} \tag{4.24}$$

whereas the linear-quadratic approximation of the problem yields the strategies

$$\begin{aligned} u_1^{lq} &= - [6x_1, 2x_2]^\top, \\ u_2^{lq}(u_1^{lq}) &= - [2x_1, 2x_2]^\top - \theta u_1^{lq}. \end{aligned} \tag{4.25}$$

Simulations have been run for the system with different combinations of the three control strategies for 25 initial conditions, $x_0 = [x_{1,0}, x_{2,0}]^\top$. These initial conditions are such that

²For clarity of presentation the dynamics of ξ is not explicitly stated in the dynamic strategies, *i.e.* u_1^d is used to denote the tuple (u_1^d, ξ) and similarly $u_2^d(u_1)$ denotes $(u_2^d(u_1), \xi)$.

they form a uniform square grid in the positive orthant, with $x_{1,0}$ and $x_{2,0}$ ranging from 0 to 4. Different values of the parameters k , R_1 , R_2 and $\xi(0)$ have been selected for the different initial conditions.

To compare the dynamic strategies resulting from (4.19) with the linear strategies (4.25) consider the quantities

$$C_1(x_0) = \frac{J_1(u_1^d, u_2^*(u_1^d)) - J_1(u_1^l, u_2^*(u_1^{lq}))}{|J_1(u_1^*, u_2^*(u_1^*))|}, \quad \text{if } J_1(u_1^*, u_2^*(u_1^*)) \neq 0, \quad (4.26)$$

$$C_2(x_0) = \frac{J_2(u_1^d, u_2^*(u_1^d)) - J_2(u_1^l, u_2^*(u_1^{lq}))}{|J_2(u_1^*, u_2^*(u_1^*))|}, \quad \text{if } J_2(u_1^*, u_2^*(u_1^*)) \neq 0.$$

Since $J_1(u_1^*, u_2^*(u_1^*)) \leq J_1(u_1^\bullet, u_2^*(u_1^\bullet))$ and $J_2(u_1^*, u_2^*(u_1^*)) \leq J_2(u_1^*, u_2^\bullet(u_1^*))$, (4.26) quantifies the loss suffered by each player when either of the players deviate from its Stackelberg strategy to the dynamic or linear strategy. More precisely, $C_1(x_0) < 0$ indicates that player 1 loses less by deviating from u_1^* to u_1^d than it does by deviating from u_1^* to u_1^{lq} . Similarly $C_2(x_0) < 0$ indicates that player 2 loses less by deviating from $u_2^*(u_1^*)$ to $u_1^d(u_1^*)$ than it does by deviating from $u_2^*(u_1^*)$ to $u_2^{lq}(u_1^*)$.

Figures 4.1 and 4.2 show the quantities $C_1(x_0)$ and $C_2(x_0)$, respectively, for the different initial conditions. Close to the origin and along the line $x_{2,0} = 0$ the linear strategies perform better than the dynamic ones. However, both $C_1(x_0)$ and $C_2(x_0)$ are small in this region. Further from the origin and the $x_{2,0} = 0$ line, along which (4.25) and (4.24) are identical, $C_1(x_0) < 0$ and $C_2(x_0) < 0$. These observations indicate that the dynamic strategies offer a good approximation of the Stackelberg equilibrium strategies in such a region. In Figures 4.1 and 4.2 dark shades indicate small values of C_i , for $i = 1, 2$, *i.e.* the dynamic strategies perform *better* than the linear strategies for player i , whereas light shades indicate the opposite.

Figure 4.3 shows the trajectories of the state x for the three different strategy pairs, namely the dynamic strategies $\mathcal{S}^d = \{u_1^d, u_2^d(u_1^d)\}$ (solid line), the linear strategies $\mathcal{S}^l = \{u_1^{lq}, u_2^{lq}(u_1^{lq})\}$ (dashed line) and the Stackelberg equilibrium strategies $\mathcal{S}^s = \{u_1^*, u_2^*(u_1^*)\}$ (dotted line) for the initial condition $x_0 = [4, 4]^\top$, which is indicated by solid dots in

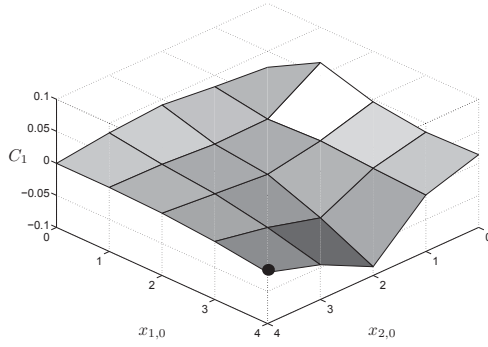


Figure 4.1: The value of $C_1(x_0)$ for various initial conditions.

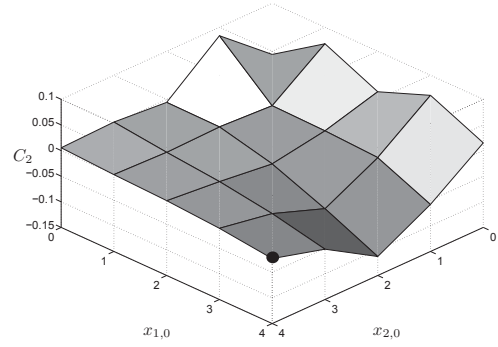


Figure 4.2: The value of $C_2(x_0)$ for various initial conditions.

Figures 4.1 and 4.2. In Figure 4.3 the square and circular markers denote the initial and the final value of the state, respectively. Table 4.5 shows the outcomes J_1 and J_2 corresponding to the strategies S^s , S^l and S^d for the same initial condition. Note that both outcomes corresponding to the players adhering to the dynamic strategies are closer to the Stackelberg equilibrium outcomes than those corresponding to the players adopting the linear strategies. Furthermore, the trajectory of the state corresponding to the players adopting the dynamic strategies is significantly closer to the trajectory corresponding to the players adhering to the Stackelberg equilibrium strategies than the trajectory resulting from the players adopting the linear strategies. Finally note that both $C_1(x_0) < 0$ and $C_2(x_0) < 0$ for this initial condition. These observations indicate that for the initial condition $x_0 = [4, 4]^\top$ the dynamic strategies provide a better approximation of the Stackelberg equilibrium solution than the linear strategies do.

	S^s	S^l	S^d
J_1	16.1100	16.4576	16.0721
J_2	16.0546	16.2284	16.0354

Table 4.1: The outcomes J_1 and J_2 corresponding to different strategies for the initial condition $x_0 = [4, 4]$.

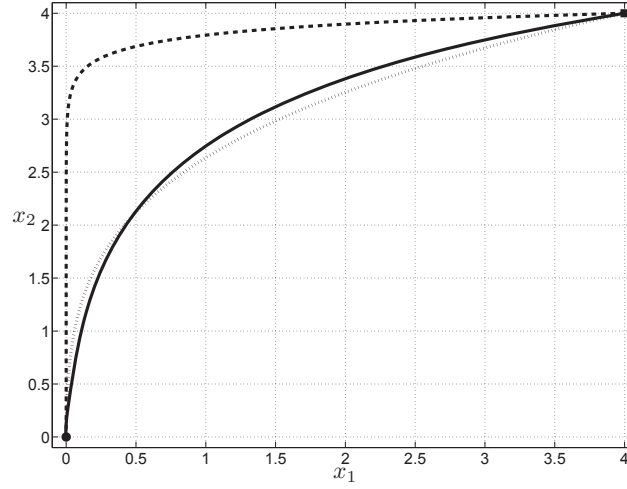


Figure 4.3: The trajectories of the state x corresponding to the control strategies S^d (solid line), S^l (dashed line) and S^s (dotted line). The square and circular markers denote the initial and final values of the state, respectively.

4.6 Conclusion

Feedback Stackelberg equilibrium solutions for a class of nonlinear nonzero-sum differential games with two players have been studied. The notion of ϵ_α -Stackelberg equilibria and *algebraic $\bar{P}$ matrix Stackelberg solution* have been introduced. Similarly to what has been done in Section 3.6, these notions are used to construct a method for designing dynamic feedback strategies that satisfy partial differential inequalities in place of the partial differential equations associated with the differential game. These strategies constitute local ϵ_α -Stackelberg equilibrium solutions for such games. The theory is illustrated by a simple numerical example for which the Stackelberg equilibrium strategies are known and it is demonstrated that the dynamic approximate solution yields a better approximation than the linear-quadratic approximation of the problem.

Chapter 5

Optimal Monitoring

5.1 Introduction

Multi-agent systems have a wide range of applications, such as monitoring and surveillance for search and rescue missions. In this chapter we consider the problem of continuously monitoring a region using a team of mobile agents equipped with sensors. These could, for example, be UAVs equipped with cameras. It is also shown that a game theoretic framework is useful for the formulation of problems involving multi-agent systems, as discussed in Section 5.3.

The problem of continuously monitoring a region using a team of agents is formulated as a differential game. One of the advantages of this problem formulation lies in that no assumptions are made on the agents, thus allowing for *heterogeneous agents*, in contrast to other approaches in which the agents are assumed to be homogeneous. The differential games considered in this chapter are somewhat different in nature to the ones studied in Chapter 3. Therefore, the framework provided by differential game theory, considered in Chapters 2 and 3, is used to formulate the problem, before an ad hoc solution to the differential game is found. More specifically, the differential game is approximated as a sequence of optimal control problems. This approach is somewhat similar to what is done in model predictive control (MPC), see for example [93,94].

The solution of each optimal control problem relies on the solution of its cor-

responding Hamilton-Jacobi-Bellman PDE. When the monitoring is performed by one agent with single-integrator dynamics, the Hamilton-Jacobi-Bellman PDEs associated with the optimal control problems can be solved analytically. However, when the monitoring task is performed by more than one agent this is not the case, *i.e.* closed-form solutions to the Hamilton-Jacobi-Bellman PDE are not available, and it is necessary to solve the optimal control problems approximately, even when the agents have simple, single-integrator dynamics. In this chapter approximate solutions are found for the case in which there are $N \geq 1$ agents. The method of obtaining approximate solutions has been discussed in [63] and is the precursor to the methods developed for differential games in Chapter 3. It relies on the use of dynamic feedback and an algebraic $\bar{P}$ solution. To solve the problem of monitoring a region a general form for algebraic $\bar{P}$ matrix solution¹ is presented, assuming the agents have single-integrator dynamics, and used to obtain approximate solutions for each optimal control problem. Furthermore, the problem of ensuring no collisions between agents occur is considered directly in the problem formulation.

Real UAVs do not satisfy single-integrator dynamics. A more accurate model of UAV dynamics is given by a unicycle model, see for example [95, 96]. The proposed method of optimal monitoring is general and could be used with different dynamics. However, for the unicycle case, a positive definite solution to the algebraic Riccati equations associated with the problem does not exist. Therefore the solution of the optimal control problem solved assuming single-integrator dynamics is used as a *trajectory plan* and a method of ensuring the UAVs with unicycle dynamics track the trajectories is proposed.

The remainder of the chapter is structured as follows. A literature review on this topic is given in Section 5.2, before focusing on approaches that make use of the framework provided by game theory. The monitoring problem is then formulated as a differential game and approximated as a sequence of infinite-horizon optimal control problems in

¹Note that we use the terminology *algebraic $\bar{P}$ matrix solution* instead of *algebraic $\bar{P}$ solution* as in [63]. This is because a definition of such a solution, which involves matrix-valued mappings, different from that used in [63] is used herein.

Section 5.4. The problem of ensuring collision avoidance is included in the formulation of each optimal control problem at this stage. To motivate the problem formulation a special case in which one agent monitors a one dimensional segment is considered in Section 5.5. Returning to the general problem, a method for solving each of the optimal control problems approximately, based on the the results presented in [63], is presented in Section 5.6. The case in which the agents have more general dynamics is then considered in Section 5.7. Simulations illustrating the results of Sections 5.6 and 5.7 are presented in Section 5.8. Finally, some concluding remarks are given in Section 5.9.

5.2 Multi-Agent Systems and Monitoring Tasks

Consider the problem of continuously monitoring a region using a team of autonomous agents equipped with sensors, similar to what has been considered in [21, 87, 97]. The team of agents could, for example, be unmanned aerial vehicles (UAVs) with on-board cameras. The task of the agents is to cooperatively monitor an area. The data gathered by the agents could provide useful information increasing both efficiency and safety during a search and rescue mission.

Various approaches for solving problems of coverage and monitoring have been studied in the literature. Most of the problems that have already been studied consider either the case in which a large number of agents survey a relatively small area, so that the problem becomes one of determining the optimal static sensor locations, or the case in which a few agents search a large area more or less randomly. In scenarios where the region to be monitored is relatively small and there exists static sensor configurations from which the entire region can be observed, the problem can be solved using localisational optimisation, where the optimal sensor placement achieving maximum coverage is sought [21, 98]. When the region is large, on the other hand, dynamic sensors are needed to cover the entire area.

The problem which we consider deals with the situation in which there does not necessarily exist a static configuration that guarantees full coverage of the region, which

is also referred to as the search space, and the number of agents is sufficiently large for a random approach to result in sub-optimal performance. It is then necessary to determine the optimal trajectories of the agents while ensuring collision avoidance and respecting constraints imposed by the agent dynamics. Similar problems have been studied in the literature, some of which are brought to light here.

The authors of [99] solve the problem of determining the best strategies to search an area using a game theoretic approach based on Nash equilibria. In particular, the search space is divided into identical hexagonal cells, each of which is associated with an uncertainty level, which together form an uncertainty map over the search space. An agent can either stay in its current cell or move to one of the neighbouring cells. The uncertainty level of a cell is reduced when an agent spends time in it and the aim is to minimise the uncertainty across the entire search space. Consequently the future actions of the agents are determined to maximise their uncertainty reduction, i.e. the agents seek to maximise their individual payoffs. These are expressed by means of search effectiveness matrices which contain all possible payoffs an agent can obtain when moving from one cell to another and a strategy is then found by looking for the non-cooperative Nash equilibrium of the matrix game defined by the search effectiveness matrices of all the participating agents. For the case in which there are two agents this becomes a sequential bimatrix games for which solutions and simulations have been presented in [99].

In [87,88] a similar problem is solved by specifying a desired coverage level across a particular planar region and a gradient-type kinematic control is developed to ensure coverage and guarantee collision-avoidance between agents. In particular the authors consider an unknown domain and define a minimum level of effective coverage. They then define an effective coverage function which quantifies how well each point of the search space has been covered over the search period. With the aim being to guarantee that the minimum effective coverage is reached across the entire search space, an error function satisfying certain conditions is defined and used to formulate a Lyapunov-like function which guarantees the convergence of a proposed control law. After discussing the need for a symmetry-breaking controller, the method is extended to include collision

avoidance.

Various other approaches are present in the literature. Some methods and examples have been presented in [100], which includes a discussion of consensus or distributed agreement. Another topic that has received a fair amount of attention recently is known as simultaneous localisation and mapping (SLAM), see for example [101, 102] and references therein.

Note that in [87, 88] there is no guarantee of optimality. In addition, some of the methods are fairly computationally expensive and result in a nonuniform coverage across the search space. Additionally, most of the coverage problems considered in the literature assume that the agents are homogeneous. In many scenarios the use of heterogeneous agents may be useful. It may, for example, be of interest to make use of both aerial and ground vehicles. In the case of a system of systems, it is evident that the individual systems are not, generally speaking, homogeneous. Thus, there is still room for further developments and different ways of addressing the problem.

Although our work is motivated by its application to search and rescue missions, several possible applications exist. In essence the problem is one of using a team of mobile sensors (which may be heterogeneous) to observe or explore some environment. In [103] drones with cameras are used for polar exploration. Similar systems can be used for structural or habitat monitoring [104, 105]. The latter could be of use for conservation projects. By monitoring certain regions useful information regarding an area and its biodiversity can be used to protect, possibly endangered, wildlife as discussed in [105], where the term *conservation drones* is used to describe the “agents”. Thus, continuously monitoring an area, in addition to being theoretically interesting, is a problem that could have a significant social impact.

5.3 Multi-Agent problems Formulated as Differential Games

In Chapters 2, 3 and 4 differential games and their solutions in terms of feedback Nash and Stackelberg equilibria have been considered. In Section 3.8.3 the *multi-agent collision*

avoidance problem is posed as a differential game and approximate solutions are obtained using the theory developed in Section 3.6. Thus, it has already been seen that differential games may be useful when formulating problems involving several agents. In this section the application of game theoretic tools to problems involving multi-agent systems (or a system of systems) is discussed further.

Recall that differential games is the study of problems in which there are several players each attempting to optimise its own cost functional and in the case of nonzero-sum differential games the cost functionals may be such that the players are either collaborating or competing. Typically multi-agent systems (or systems of systems) are such that agents (or individual systems) solve a task collaboratively, although there may also be some competition between agents as seen in [106]. Simply replacing the word “agent” with “player” in the term *multi-agent systems*, it appears natural that the differential game framework may be well-suited to deal with problems involving multi-agent systems. Namely, each agent in a multi-agent system can be considered as a player and their tasks can be described by suitable cost functionals which they seek to optimise subject to dynamics which describe their collective behaviour. The multi-agent collision avoidance problem in Section 3.8.3 is one example of this.

There are other examples available in the literature where game theoretical approaches have been taken to solve problems involving multi-agent systems. Two examples of this, which have already been mentioned, are the problems considered in [106] and [99]. In [106] a team of agents pursue a team of evaders in an unknown environment and the problem is formulated as a probabilistic pursuit-evasion game, whereas in [99] the problem of covering an unknown region is formulated as a sequence of bimatrix games. There are several other examples of problems involving a multi-agent system or a system of systems that suggest that the framework provided by differential games may provide a suitable way of formulating and solving the problems. In [25] the problem of detecting a threat in an unknown environment, known as *security monitoring* is considered. The task is performed by three systems, namely a master robot, a sensor network and a swarm of robots, which together form a system of systems. These heterogeneous

systems seek to achieve a common goal, however the requirements for each system are not identical and may be conflicting. The objectives of each system are described by appropriate cost functionals and a linear combination of these is considered as an overall cost functional, resulting in a (static) optimisation problem. However, the nature of the problem suggests that an alternative approach may be possible using the tools provided by a game theoretical framework. Similarly, in [107] the problem of planning the paths of mobile robots in an environment with sensor nodes subject to different objectives, such as minimising power consumption and maximising data collection, is studied. The problem is solved using a fuzzy logic approach to solve discrete decision problems, referred to as multi-criteria decision making. An alternative (continuous-time) problem formulation could be possible by posing the problem as a differential game.

5.4 Problem Formulation

We now return our focus to the problem of optimally monitoring a region using a team of agents.

Consider the problem of monitoring a region, $\Omega \subset \mathbb{R}^n$, where $n > 0$, using N agents, each equipped with a sensor. Let $q_i(t) \in \mathbb{R}^n$ denote the position of agent i at time t and suppose its sensor can be modeled by a function, $S_i(q_i(t), q) \geq 0$, where $q \in \Omega$ and $i = 1, \dots, N$. In what follows, unless otherwise stated, it is assumed that $n = 2$, *i.e.* the region to be monitored is two dimensional. The *coverage map*

$$J(q, T) = \int_0^T \sum_{i=1}^N S_i(q_i(t) - q) dt, \quad (5.1)$$

quantifies the coverage level of a point $q \in \Omega$ after the agents have been monitoring the area over the period $[0, T]$ along $q_i(t)$, $i = 1, \dots, N$. This function is equivalent to the *coverage function* considered in [87, 88].

Assumption 9. The sensor model of agent i is locally quadratic, *i.e.*

$$S(q_i - q) = A_i - (q_i - q)^\top \bar{Q}_i (q_i - q) + \text{h.o.t.},$$

where $q \in \Omega$, $A_i > 0$ and the matrix $\bar{Q}_i \in \mathbb{R}^{2 \times 2}$ is positive definite. In broad terms this implies that the coverage of agent i has a maximum at the agent's position, q_i , and exhibits a quadratic decay close to this point.

In what follows it is assumed that Assumption 9 is satisfied for $i = 1, \dots, N$, *i.e.* all agents are equipped with sensors with models that are at least locally quadratic. Note, however, that the sensor models $S_i(q_i, q)$ need not be identical for all $i = 1, \dots, N$. This allows for the use of *heterogeneous* agents, differently from [87, 88].

Remark 21. Figure 5.1 shows the coverage provided by three different sensor models satisfying Assumption 9 for an agent positioned at $q_i = [0, 0]^\top$. In particular Figure 5.1 (a) shows a locally quadratic model given by

$$S^1(q_i - q) = \begin{cases} \frac{1}{2^4} (\|q_i - q\|^2 - 4) & \text{if } \|q_i - q\| \geq 2, \\ 0 & \text{otherwise,} \end{cases} \quad (5.2)$$

similar to what has been used in [87, 88], whereas (b) shows an exponential sensor model as used in [41], *i.e.*

$$S^2(q_i - q) = e^{-0.8(q_i - q)^2}, \quad (5.3)$$

and (c) shows a model given by

$$S^3(q_i - q) = 1 - \frac{5\|q_1 - q\|^2}{1 + 5\|q_1 - q\|^2}. \quad (5.4)$$

In Figure 5.1 the plots of the sensor models have been normalised so that the peak coverage is 1 for all three cases (dark colours indicate a low level of coverage). ▲

Note that $S^1(q_i - q)$ has compact support, whereas the other two models do not.

Intuitively, monitoring the region Ω can be achieved by continuously attempting to maximise the coverage of the points $q \in \Omega$ where the coverage is low. With this in mind, *virtual players*, $\tilde{q}_i$, $i = 1, \dots, N$, with *relatively fast dynamics* are introduced. One virtual player is introduced for each agent and the idea behind this is that the virtual players should “identify” points where the coverage is low. It is assumed that the agents

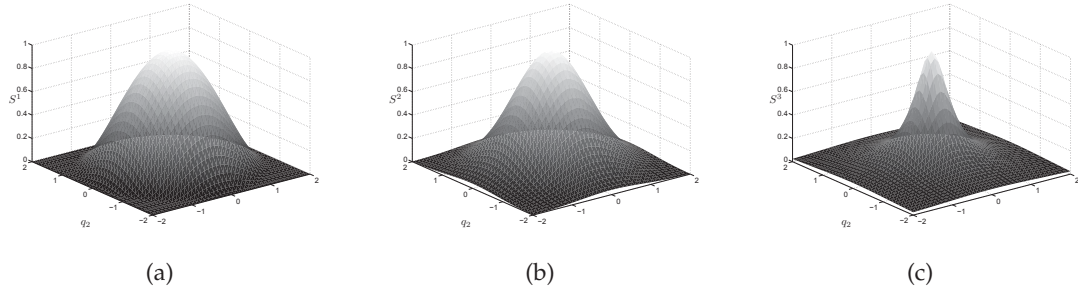


Figure 5.1: Three sensor models (normalised such that the peak value is 1) satisfying Assumption 9: (a) $S^1(q)$, (b) $S^2(q)$ and (c) $S^3(q)$.

satisfy the single-integrator dynamics

$$\dot{q}_i = u_i, \quad (5.5)$$

with $q_i(t) \in \mathbb{R}^2$ the state of agent i and $u_i(t) \in \mathbb{R}^2$ the control action of agent i . Similarly, the virtual players are attributed the dynamics²

$$\dot{\tilde{q}}_i = w_i, \quad (5.6)$$

with $\tilde{q}_i \in \mathbb{R}^2$, $i = 1, \dots, N$, the states of the virtual players. Furthermore, the control efforts of the agents are such that $\|u_i\|^2 \leq U$ and $\|w_i\|^2 \leq W$, where $W \gg U$.

Remark 22. Although it is assumed here that the dynamics of all agents are subject to the same constraints, different bounds can easily be attributed to each agent and the same holds for the virtual players. ▲

The problem of continuously monitoring a region can then be formulated as follows.

Problem 11. The problem of monitoring the region consists in determining strategies u_i and w_i , for $i = 1, \dots, N$, such that the set of strategies $\mathcal{S} = \{u_1, w_1, \dots, u_N, w_N\}$ constitutes a Nash equilibrium strategy for differential game with cost functional³

$$\min_{u_1, \dots, u_N} \max_{w_1, \dots, w_N} -\frac{1}{2} \int_0^T \sum_{i=1}^N S_i(q_i(t) - \tilde{q}_i(t)) dt, \quad (5.7)$$

²Since these players are virtual, they can be attributed arbitrary dynamics.

³Note that the integral in (5.7) is the coverage map (5.1).

and constrained dynamics (5.5) and (5.6), $i = 1, \dots, N$. Furthermore the strategies should be such that $q_i \in \Omega$ and $\tilde{q}_i \in \Omega$ for all $i = 1, \dots, N$.

The cost functional (5.7) is such that the virtual players seek to move towards regions with low coverage, whereas each agent seeks to minimise the distance between itself and its corresponding virtual player.

Remark 23. Problem 11 is somewhat different in nature from the differential games considered in Chapter 3. One of the main differences is that in Problems 3 and 4 infinite horizon differential games are considered and these are such that all players “agree” on the equilibrium of the dynamic system, namely the cost functionals are such that $q_i(0) = 0$ for all $i = 1, \dots, N$ and as a consequence the players drive the states to the zero-equilibrium while attempting to minimise their individual cost functionals. This is not the case in Problem 11 in which the players seek to reach different positions within the region Ω . Additionally the differential game in Problem 11 does not terminate as the monitoring should be continuous. Extensions of the results in Chapter 3 for finite-horizon differential games may be useful to deal with differential games of this kind. ▲

To solve the differential game described by equations (5.5), (5.6), $i = 1 \dots, N$, and (5.7), the fast dynamics of the virtual players are taken advantage of. In particular it is assumed that the virtual players dynamics are sufficiently fast such that they can be approximated as *instantaneous players*. The interval $[0, T]$ is split into smaller intervals of length τ , such that each of the optimal control problems is taken to be of duration τ and the agents seek to minimise the cost functional

$$\frac{1}{2} \int_{t_0}^{t_0+\tau} \left(- \sum_{i=1}^N S_i(q_i - \tilde{q}_i) + \sum_{i=1}^N \alpha_i \|u_i\|^2 \right) dt, \quad (5.8)$$

where $\alpha_i > 0$ are weights penalising the control efforts of the agents and t_0 denotes the beginning of each optimal control problem. The *instantaneous players* are updated between successive optimal control problems. Note that the constraints on u_i are translated into soft constraints by adding a term to the cost functional which penalises the control effort. Different bounds on the control efforts can be “enforced” by tuning α_i .

Provided τ is sufficiently large⁴, each of the optimal control problems defined by the cost functional (5.8) and the agent dynamics (5.5) can be further approximated to have an infinite-horizon. Each of the infinite-horizon optimal control problems is then characterised by the agent dynamics (5.5) and the cost functional which each of the agents attempts to minimise, *i.e.*

$$\frac{1}{2} \int_0^\infty \left(C - \sum_{i=1}^N S_i(q_i - \tilde{q}_i) + \sum_{i=1}^N \alpha_i \|u_i\|^2 \right) dt, \quad (5.9)$$

where $C = \sum_{i=1}^N A_i \geq 0$ is a constant ensuring the integrand is zero at the minimiser of the integrand, namely when $q_i = \tilde{q}_i$, for $i = 1, \dots, N$.

Note that when $N > 1$, additional care must be taken to ensure that inter-agent collisions do not occur. In particular, suppose each agent i , for $i = 1, \dots, N$, seeks to maintain a minimum *safety distance* $r_i > 0$ between itself and the other agents and that collisions are only considered a risk when the distance between itself and another agent is below a *risk distance*, $R_i > r_i$. Given R_i and r_i , for $i = 1, \dots, N$, suppose the following assumptions, similar to Assumptions 5 and 6 introduced for the *multi-agent collision avoidance problem*, are satisfied.

Assumption 10. The initial conditions $q_1(0), \dots, q_N(0)$ are such that $\|q_i(0) - q_j(0)\|^2 > R_i^2$ for all $i = 1, \dots, N, j = 1, \dots, N$ and $j \neq i$.

Assumption 11. The values taken by the virtual players, namely the $\tilde{q}_1, \dots, \tilde{q}_N$, are such that $\|\tilde{q}_i - \tilde{q}_j\|^2 > R_i^2$ for all $i = 1, \dots, N, j = 1, \dots, N$ and $j \neq i$.

Let $q = [q_1^\top, \dots, q_N^\top]^\top \in \mathbb{R}^{2N}$, $\tilde{q} = [\tilde{q}_1^\top, \dots, \tilde{q}_N^\top]^\top \in \mathbb{R}^{2N}$ and $x = q - \tilde{q}$, and define the *collision avoidance function*

$$v_i(q) = \left(\min \left(0, \sum_{j=1, j \neq i}^N \frac{\|q_i - q_j\|^2 - R_i^2}{\|q_i - q_j\|^2 - r_i^2} \right) \right)^2. \quad (5.10)$$

In [87, 88] this function has been used as a barrier Lyapunov functions to ensure collision

⁴In this case, τ sufficiently large means that the agents have enough time to reach the "minimisers" of (5.8).

avoidance. Herein, similarly to what has been done in Section 3.8.3, collision avoidance is included directly in the formulation of the optimal control problem. Then the problem of collision avoidance can be dealt with by adding the *collision avoidance function* (5.10) to the integrand of (5.9) and defining the optimal control problem as follows.

Let $u = [u_1^\top, \dots, u_N^\top]^\top$ and assume without loss of generality, that $\alpha_i = 1$, for $i = 1, \dots, N$. Then each optimal control problem is defined by the cost functional

$$\frac{1}{2} \int_0^\infty q(x) + \|u\|^2 dt, \quad (5.11)$$

which the agents seek to minimise, where

$$q(x) = q_{cov}(x) + \beta q_{col}(x, \tilde{q}), \quad (5.12)$$

with $q_{cov}(x) = C - \sum_{i=1}^N S_i(q_i - \tilde{q}_i)$, $q_{col}(x) = x^\top \text{diag}\{v_1(x - \tilde{q}), \dots, v_N(x - \tilde{q})\}x$ and $\beta > 0$, and the dynamics

$$\dot{x} = u. \quad (5.13)$$

Note that if the Hamilton-Jacobi-Bellman PDE associated with the problem, namely⁵

$$-\frac{1}{2} \frac{\partial V}{\partial x} \frac{\partial V}{\partial x}^\top + \frac{1}{2} q(x) = 0, \quad (5.14)$$

with $V(0) = 0$, can be solved, the optimal control actions are given by

$$u^* = -\frac{\partial V}{\partial x}^\top \quad (5.15)$$

Remark 24. It follows from Assumption 9 that the part of the running cost (5.12) relating to the coverage can be written in the form $q_{cov}(x) = x^\top Q_{cov}(x)x$. Thus, the whole running cost can be written as $q(x) = x^\top Q(x)x$, where $Q(x) = Q_{cov}(x) + \text{diag}\{v_1(x + \tilde{q}), \dots, v_N(x + \tilde{q})\}$. In a neighbourhood of the origin the remaining cost can be approximated as⁶ $q(x) = x^\top \bar{Q}x$, where $\bar{Q} = Q(0) = \text{diag}\{\bar{Q}_1, \dots, \bar{Q}_N\}$. ▲

⁵ F_r is used to denote the gradient of $F(\cdot)$ with respect to r , i.e. $F_r = \nabla_r F$.

⁶Note that $v_i(x + \tilde{q})|_{x=0} = 0$ for $i = 1, \dots, N$, which follows from (5.10) and Assumption 11.

The virtual player, $\tilde{q}_i$, which is taken to act *instantaneously*, is updated between successive optimal control problems according to

$$\tilde{q}_i = \underset{p}{\operatorname{argmin}} J(p, t_0 + \tau), \quad (5.16)$$

where $J(p, t)$ is the coverage map introduced in (5.1), $p \in \left\{ \Omega : \|q_i - p\|^2 < R_{d,i}^2, \|\tilde{q}_i - \tilde{q}_j\|^2 > \max\{R_i, R_j\}^2 \right\}$, and $R_{d,i} \geq 0$. Loosely speaking, this means that the *virtual player*, $\tilde{q}_i$ moves to the least covered point in Ω which is within a radius $R_{d,i}$ from its current position, while ensuring the selection is such that Assumption 11 holds⁷.

Remark 25. Since the virtual players $\tilde{q}_i, i = 1, \dots, N$ have fast dynamics and are approximated to act *instantaneously* between successive optimal control problems according to (5.16), during each optimal control problem the variables $\tilde{q}_i, i = 1, \dots, N$ are constant and it therefore follows that $\dot{x} = [\dot{q}_1^\top, \dots, \dot{q}_N^\top]^\top$. The actions of the virtual players are represented by the instantaneous updates (5.16). ▲

Remark 26. Agent dynamics different from (5.5) can be considered in this problem formulation provided the corresponding Hamilton-Jacobi-Bellman PDE is appropriately modified. ▲

5.5 Special case: Single Agent on a Segment

In this section a special case in which one agent equipped with a sensor modeled by (5.3) is considered: in this case closed-form solutions the (5.14) can be found. Furthermore, we focus on the problem in which the agent monitors a one-dimensional (1D) segment, *i.e.* the Hamilton-Jacobi-Bellman PDE (5.14) boils down to a simple ordinary differential equation (ODE). The results discussed in this section are consistent with what one would intuitively expect for this simple problem, thus suggesting that the problem formulation and the approach taken to obtain an approximate solution for Problem 11 is reasonable.

⁷Note that the allocation of $\tilde{q}_1, \dots, \tilde{q}_N$ may be non-unique.

Consider the optimal monitoring problem for a single agent on a segment, *i.e.* consider the case in which $N = 1$ and $\Omega \subset \mathbb{R}$.

Proposition 4. Let $n = 1$, $N = 1$, $C = 1$ and let S_1 be given by (5.3). Consider the optimal control problem with cost functional (5.9) and agent dynamics (5.13). Then, the optimal control is given by

$$u_i^* = \begin{cases} -\frac{\sqrt{1 - e^{A_1(q_1(t) - \tilde{q}_1)^2}}}{\alpha_i} & \text{if } q_1 \geq \tilde{q}_1, \\ +\frac{\sqrt{1 - e^{A_1(q_1(t) - \tilde{q}_1)^2}}}{\alpha_i} & \text{if } q_1 \leq \tilde{q}_1. \end{cases} \quad (5.17)$$

◇

Proof: The Hamilton-Jacobi-Bellman equation (5.14) for this problem is

$$\frac{1}{2} - \frac{1}{2} e^{A_1(q_1(t) - \tilde{q}_1)^2} - \frac{V_{q_1}^2}{2\alpha_1} = 0, \quad (5.18)$$

which can be trivially integrated to yield the solution

$$V_{q_1} = \begin{cases} +\frac{\sqrt{1 - e^{A_1(q_1(t) - \tilde{q}_1)^2}}}{\alpha_1} & \text{if } q_1 \geq \tilde{q}_1, \\ -\frac{\sqrt{1 - e^{A_1(q_1(t) - \tilde{q}_1)^2}}}{\alpha_1} & \text{if } q_1 \leq \tilde{q}_1, \end{cases} \quad (5.19)$$

where the signs are chosen to ensure that V is positive definite. The optimal control follows from (5.15). □

The value taken by the virtual player, *i.e.* $\tilde{q}_1$ is updated between successive optimal control problems according to (5.16).

To illustrate the results, consider a 1D region defined as $\Omega = \{q : -10 \leq q \leq 10\}$. Simulations have been run for 20 iterations with $A_1 = 1$, $\tau = 30$, $\alpha_1 = 1$, $q_1(0) = -5$, $R_{d,1} = 20$ and the first choice of $\tilde{q}_1 = -5$. Figure 5.2 shows the time history of the position of the agent, q_1 . The coverage levels at each point, as given by (5.1), are initially zero and their time-evolution, along with the evolution of the static minimiser, $\tilde{q}$, is shown in Figure 5.3. It can be observed that the minimiser moves towards regions that have relatively low coverage levels. The minimum and maximum coverage levels across the

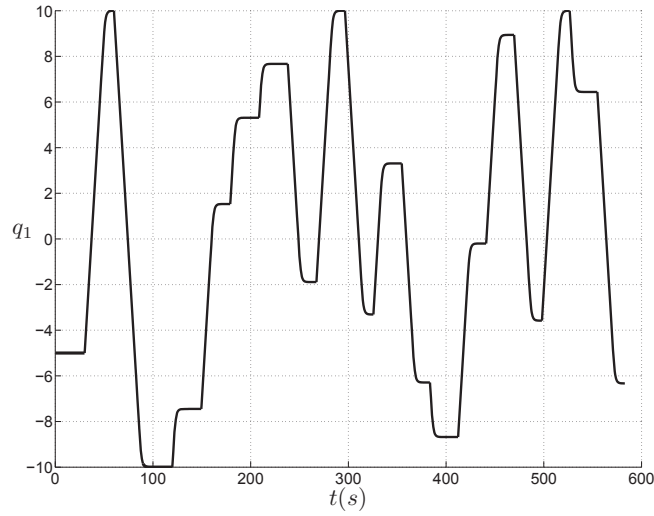


Figure 5.2: The time history of the position of the agent, q_1 , with $q_1(0) = -5$ and $\tilde{q}_1(0) = -5$ for $t \in [0, 600]$.

entire region are shown as a function of time in Figure 5.4, which shows that both values are monotonically increasing with time, thus indicating that the information gathered by the agent is increasing with time, *i.e.* the agent is indeed continuously monitoring the search space. Since the simulations have been run by solving the optimal control problems over a fixed time horizon, the time parametrization is not significant and it should be noted that the overall behaviour is a sweeping motion as one would expect.

5.6 Approximate Solutions to the Optimal Control Problems

In the special case considered in Section 5.5 closed-form solutions to equation (5.14) are found and the simulations show that the results are consistent with expectations. We now focus our attention on the general problem discussed in Section 5.4. In general, closed-form solutions to the Hamilton-Jacobi-Bellman PDEs associated with the optimal control problem defined by (5.9) and the agent dynamics cannot be found. It is therefore necessary to seek approximate solutions. This can be done using the method presented in [37, 63], namely by introducing a dynamic extension and using the notion of *algebraic $\bar{P}$ matrix solution*.

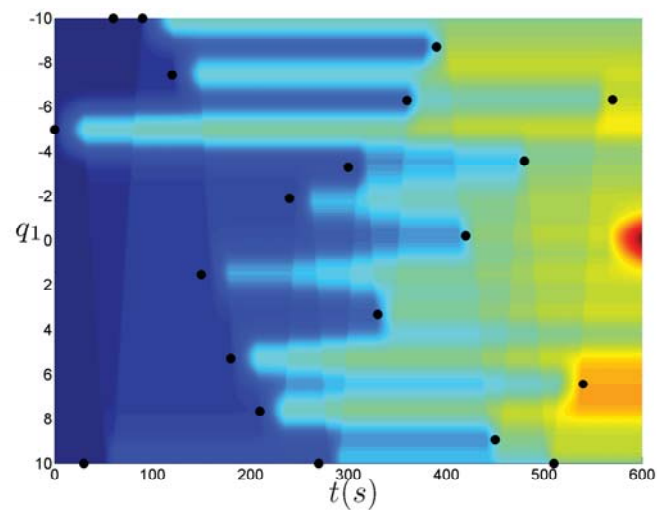


Figure 5.3: Values of $\tilde{q}_1$ (solid dots). The colour map illustrates the coverage levels of Ω at each time (blue: low coverage level, red: high coverage level).

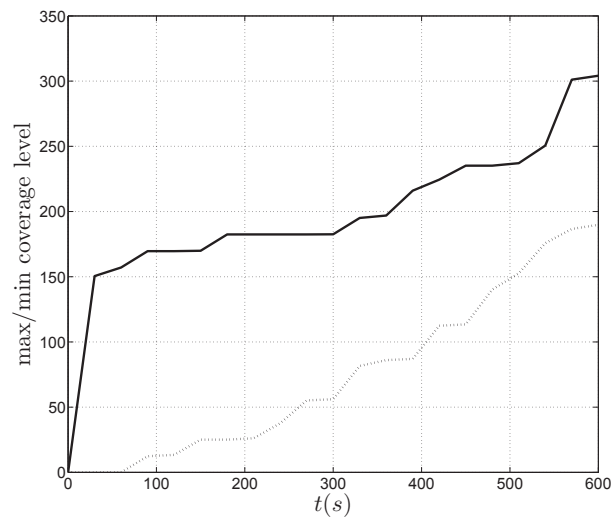


Figure 5.4: Time history of minimum (dotted line) and maximum (solid line) coverage levels in Ω .

To streamline the presentation of the following proposition, the main results of [63] are specialised to the current problem. Let $\sigma(x) = x^\top \Sigma(x)x \geq 0$. A mapping $P(x) = P(x)^\top \in \mathbb{R}^2$ satisfying

$$-P(x)P(x) + Q(x) + \Sigma(x) = 0, \quad (5.20)$$

where $Q(x)$ is as defined in Remark 25, and $P(0) = \bar{P}$ is the positive definite solution of the continuous time ARE

$$-\bar{P}\bar{P} + \bar{Q} + \bar{\Sigma} = 0, \quad (5.21)$$

where $\bar{\Sigma} = \Sigma(0)$, is said to be an *algebraic $\bar{P}$ matrix solution* of the optimal control problem with cost functional (5.11) and dynamics (5.13). Let $P(x)$ denote such a solution, introduce a dynamic extension $\xi \in \mathbb{R}^{2N}$ and define

$$V(x, \xi) = \frac{1}{2}x^\top P(\xi)x + \frac{1}{2}\|x - \xi\|_R^2, \quad (5.22)$$

where $R = R^\top > 0 \in \mathbb{R}^{2N \times 2N}$. It can then be shown that the dynamic strategies

$$\begin{aligned} u(x, \xi) &= -\frac{\partial V}{\partial x}^\top \\ \dot{\xi}(x, \xi) &= -k \frac{\partial V}{\partial \xi}^\top, \end{aligned} \quad (5.23)$$

provide an “approximate” solution of the optimal control problem [63]. More precisely, the dynamic control (5.23) is such that the partial differential inequality

$$-\frac{1}{2} \frac{\partial V}{\partial x} \frac{\partial V}{\partial x}^\top + \frac{1}{2} q(x) + \frac{\partial V}{\partial \xi} \dot{\xi} \leq 0, \quad (5.24)$$

is satisfied along closed-loop trajectories of the system. This implies that the control (5.23) is the solution of a modified optimal control problem with dynamics (5.13) and cost functional

$$\frac{1}{2} \int_0^\infty q(x) + \|u\|^2 + c(x, \xi) dt,$$

where the function $c(x, \xi) \geq 0$ can be interpreted as a *confidence interval* and $\frac{1}{2} \int_0^\infty c(x, \xi) dt$ is the cost incurred by the approximation.

With this in mind, consider the extended system with state $[x^\top, \xi^\top]^\top$, with $\xi \in$

$\mathbb{R}^{2N}$. In particular, let $\xi = [\xi_1^\top, \dots, \xi_N^\top]^\top$ where $\xi_i \in \mathbb{R}^2$ for $i = 1, \dots, N$. Let $\Sigma(x) = \beta q_{col}(x, \tilde{q}) + \gamma I$, where $\beta > 0$ and $\gamma \geq 0$. Suppose $\bar{P}$ is the symmetric, positive definite solution to the ARE corresponding to the linear-quadratic approximation of the infinite-horizon optimal control problem with an additional running cost $\sigma(x)$, namely $\bar{P}$ is the positive definite solution of the ARE (5.21), where $\bar{Q}$ is as defined in Remark 24 and Assumption 11 implies $q_{col}(0, \tilde{q}) = 0$ and therefore $\bar{\Sigma} = \gamma I$. Suppose we can find a matrix-valued function, $\Delta(x) = \Delta(x)^\top$, such that

$$\bar{P}\Delta^\top(x)\Delta(x)\bar{P} = Q(x) + \Sigma(x), \quad (5.25)$$

where $Q(x)$ is as defined in Remark 24, and

$$\Delta(0) = I. \quad (5.26)$$

Define the matrix-valued function

$$P(x) = \bar{P}\Delta(x)^\top, \quad (5.27)$$

and let $\Psi(\xi, x)$ denote the Jacobian matrix of $\frac{1}{2}P(\xi)x$ with respect to ξ . Then the following proposition holds.

Proposition 5. There exists $k > 0$, $R = R^\top > 0$ and $\xi(0)$ satisfying $\|\xi_i(0) + \tilde{q}_i\|^2 > r_i^2$, $i = 1, \dots, N$, such that the dynamic control law

$$\begin{aligned} u &= -(\Delta(\xi)\bar{P}x + R(x - \xi)), \\ \dot{\xi} &= -k(\Psi(\xi, x)^\top x - R(x - \xi)), \end{aligned} \quad (5.28)$$

solves the optimal control problem described by (5.11), (5.12) and (5.13) approximately, in the sense that the partial differential inequality (5.24) is satisfied, and guarantees that

$$\lim_{t \rightarrow +\infty} q_i(t) = \tilde{q}_i, \quad (5.29)$$

for $i = 1, \dots, N$, i.e. each agent reaches the position of the corresponding virtual player.

Furthermore, $\|q_i(\bar{t}) - q_j(\bar{t})\|^2 > (\max\{r_i, r_j\})^2$ for all $\bar{t} > 0, i = 1, \dots, N, j = 1, \dots, N, j \neq i$. $\diamond$

Proof: The proof boils down to showing that (5.27) is an *algebraic $\bar{P}$ matrix solution* of the Hamilton-Jacobi-Bellman equation (5.14). The rest follows from [63].

Note that $P(0) = \bar{P}$ follows directly from (5.27) and (5.26) and a direct substitution of $x = 0$ in (5.25) shows that (5.21) is satisfied. Furthermore (5.20) is implied by (5.25). This in turn implies that $P(x)$ is indeed an *algebraic $\bar{P}$ matrix solution* of (5.14). The dynamic controller (5.28) is given by equations (5.23) and it follows from [63] that (5.28) solves the partial differential inequality (5.24) and thus solves the optimal control problem approximately. Thus, the origin of the extended state space $[x^\top, \xi^\top]^\top$ is asymptotically stable. The latter implies that (5.29) is satisfied for $i = 1, \dots, N$. In addition the function (5.22) is such that $V(x(0), \xi(0)) = \frac{1}{2} \int_0^\infty q(x) + \|u\|^2 + c(x, \xi) dt$, where $c(x, \xi) \geq 0$, and $\dot{V}(x, \xi) \leq -\frac{1}{2}q(x) \leq 0$. By Assumption 10 and since $\|\xi_i(0) + \tilde{q}_i\|^2 > r_i^2$ for $i = 1, \dots, N$, it follows that $V(x(0), \xi(0))$ is bounded and by Assumption 11 it follows that the optimal control problem is well-posed. Then, it follows from standard Lyapunov arguments that $\|q_i(\bar{t}) - q_j(\bar{t})\|^2 > (\max\{r_i, r_j\})^2$ for all $\bar{t} > 0, i = 1, \dots, N, j = 1, \dots, N, j \neq i$ [38]. $\square$

At the beginning of each optimal control problem the dynamic extension ξ must be initialised. This selection of $\xi(0)$ has an impact on the performance of the approximate solution to the optimal control problem. The instantaneous players are updated according to (5.16) at the end of each optimal control problem.

5.7 Unicycle Agent Dynamics

Real UAVs do not have single-integrator dynamics. However taking into consideration more complicated dynamics closed-form solutions to the Hamilton-Jacobi-Bellman PDEs cannot generally be found. Thus, it becomes necessary to settle for approximate solutions as in the previous section. UAVs are often modeled by unicycle dynamics [95, 96]. In this case, however, not even a solution to the ARE (5.21) exists. Hence an alternative method

has to be developed. To this end the trajectories resulting from the single-integrator dynamics is interpreted as a *trajectory plan* and the actual controllers are designed to track this *trajectory plan* assuming that each agent has unicycle dynamics, namely

$$\begin{aligned}\dot{x}_i &= v_i \cos \theta_i, \\ \dot{y}_i &= v_i \sin \theta_i, \\ \dot{\theta}_i &= \omega_i,\end{aligned}\tag{5.30}$$

$i = 1, \dots, N$, where $p_i(t) = [x_i(t), y_i(t)]^\top \in \mathbb{R}^2$ denotes the position and $\theta_i(t) \in \mathbb{R}$ denotes the orientation of agent i . The variables $v_i(t) \in \mathbb{R}$ and $\omega_i(t) \in \mathbb{R}$ are the control inputs of the system determining the forward speed and the angular velocity of the agent, respectively. One of the challenges in this approach lies in that the *trajectory plan* may generate trajectories that cannot be tracked with zero error.

Assuming the actual UAVs satisfy the dynamics (5.30), it is desired that $p_i(t)$ tracks $q_i(t) = [q_{i,x}(t), q_{i,y}(t)]^\top$ given by the solution of the sequence of optimal control problems defined by (5.11) and (5.13). Let $e_i(t) = p_i(t) - q_i(t)$ and define $\rho_i^2 = (x_i - q_{i,x})^2 + (y_i - q_{i,y})^2$ and $\phi_i = \arctan(y_i - q_{i,y}, x_i - q_{i,x})$.⁸ The motivation behind the change of coordinates is that the zero equilibrium of the error variable is stabilisable in the polar representation. Since it is desired that the agents with unicycle dynamics track the *trajectory plan* it is of interest to ensure the error remains bounded and converges to zero.

The signal $\dot{q}_i$ is known from the solution of the optimal control problem (5.11), (5.13) and can be considered as an external signal or disturbance. Thus, let $\delta_i = [\dot{q}_{i,x}, \dot{q}_{i,y}]^\top$ and consider the following assumption.

Assumption 12. The signal δ_i decays exponentially and furthermore $\delta_i(t)$ converges to zero faster than $\rho_i(t)$ so that $\begin{bmatrix} \frac{-\sin(\phi_i)}{\rho_i} & \frac{\cos(\phi_i)}{\rho_i} \end{bmatrix} \delta_i(t)$ can be bounded from above by a constant $M > 0$.

Remark 27. It follows from Proposition 5 that the zero equilibrium of the closed-loop system (5.13) with controllers (5.28) is locally asymptotically stable. This in turn implies

⁸The four-quadrant version of the arctangent is used to distinguish between opposite directions.

that δ_i decays exponentially, *i.e.* Assumption 12 is reasonable. $\blacktriangle$

Proposition 6. Consider the dynamics (5.30) and the trajectories generated by the algorithm in Section 5.6. Suppose Assumption 12 is satisfied. The controller

$$\begin{aligned} v_i &= -k_v \cos(\phi_i - \theta_i) \rho_i, \\ \omega_i &= -k_\omega (\theta_i - \phi_i) + k_v \sin(\phi_i - \theta_i) \cos(\phi_i - \theta_i), \end{aligned} \quad (5.31)$$

with $k_v > 0$, $k_\omega > 0$, is such that

$$\lim_{t \rightarrow \infty} e_i(t) = 0. \quad (5.32)$$

$\diamond$

Proof: First, note that

$$\begin{bmatrix} \dot{\rho}_i \\ \dot{\phi}_i \end{bmatrix} = \begin{bmatrix} \cos(\phi_i - \theta_i) \\ -\frac{\sin(\phi_i - \theta_i)}{\rho_i} \end{bmatrix} v_i - \begin{bmatrix} \cos(\phi_i) & \sin(\phi_i) \\ -\frac{\sin(\phi_i)}{\rho_i} & \frac{\cos(\phi_i)}{\rho_i} \end{bmatrix} \delta_i. \quad (5.33)$$

Then, with v_i and ω_i as defined in (5.31),

$$\dot{\rho}_i = -k_v \cos(\phi_i - \theta_i)^2 \rho_i - \begin{bmatrix} \cos(\phi_i) & \sin(\phi_i) \end{bmatrix} \delta_i,$$

and by Assumption 12 the last term of $\dot{\phi}_i$ satisfies

$$\dot{\theta}_i - \dot{\phi}_i \leq -k_\omega \left[(\theta_i - \phi_i) - \frac{M}{k_\omega} \right].$$

It follows that $\lim_{t \rightarrow \infty} \rho_i(t) = 0$, which implies that $\lim_{t \rightarrow \infty} e_i(t) = 0$. Furthermore provided δ_i converges to zero faster than ρ_i , $\theta_i - \phi_i$ remains bounded. $\square$

Remark 28. The *trajectory plan* may provide trajectories that are not feasible for the agents with unicycle dynamics to follow. Therefore, it is not possible to guarantee that $\lim_{t \rightarrow \infty} \theta_i(t) - \phi_i(t) = 0$. Nevertheless Proposition 6 implies that the error variable converges to zero. $\blacktriangle$

Remark 29. Although collision avoidance is guaranteed for the agents with single-integrator dynamics by Proposition 5, this can no longer be guaranteed for the agents

with unicycle dynamics, since the *trajectory plan* may contain infeasible trajectories. However, an upper bound for ρ_i can be found as a function of k_v and k_ω and thus through the selection of these control parameters, and $R_1, \dots, R_N$ and $r_1, \dots, r_N$, it is possible to ensure that collisions do not occur between the agents with unicycle dynamics. $\blacktriangle$

5.8 Simulations

In this section we present simulation results for the case in which $N = 2$ and

$$\Omega = \left\{ \begin{pmatrix} x \\ y \end{pmatrix}^\top : -10 \leq x \leq 10, -10 \leq y \leq 10 \right\},$$

i.e. two agents monitor a square region centred at the origin. Suppose the agents are equipped with sensors modeled by

$$\begin{aligned} S_1(q_1 - q) &= \frac{a_1}{a_3} - \frac{a_1 \|q_1 - q\|^2}{a_2 + a_3 \|q_1 - q\|^2}, \\ S_2(q_2 - q) &= \frac{b_1}{b_3} - \frac{b_1 \|q_2 - q\|^2}{b_2 + b_3 \|q_2 - q\|^2}, \end{aligned}$$

where $a_i > 0, b_i > 0, i = 1, 2, 3$. It follows that

$$P(x) = \text{diag} \left(\sqrt{\frac{a_1}{a_2 + a_3 \|q_1 - q\|^2} + (\gamma + \beta)v_1 + \gamma I_2}, \right. \\ \left. \sqrt{\frac{b_1}{b_2 + b_3 \|q_2 - q\|^2} + (\gamma + \beta)v_2 + \gamma I_2} \right),$$

with $\gamma \geq 0$, is an *algebraic $\bar{P}$ matrix solution* of the corresponding optimal control problem.

Let $a_1 = 5, a_2 = 0.5, a_3 = 10, b_1 = 2, b_2 = 1$ and $b_3 = 10$. For the case in which the agents are assumed to have single-integrator dynamics, the dynamic controller (5.28) with $k = 2, R = 0.5I, \gamma = 20, r_1 = r_2 = r = 1, R_1 = R_2 = R = 2, \beta = 0.1, \gamma = 1$ and $\xi(0) = [-20, 20, 20, -20]^\top$ has been used. As discussed in Section 5.7, the trajectories resulting from the assumption that the agents have single-integrator dynamics are interpreted as *trajectory plans* and the controllers (5.31), $i = 1, 2$, are adopted by the agents, with $k_v = 1, k_\omega = 10$.

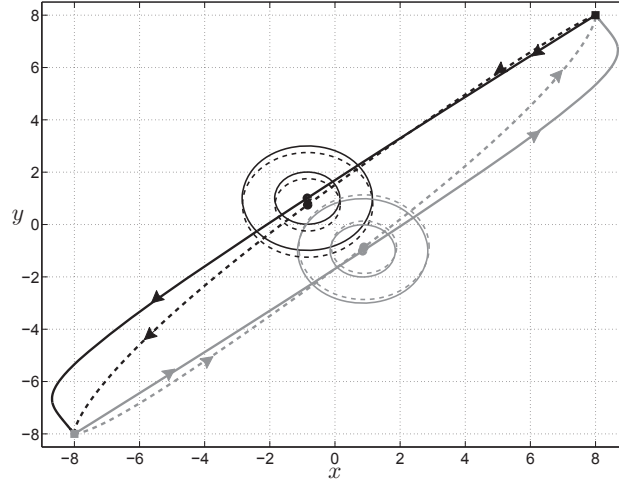


Figure 5.5: Trajectories of agents 1 and 2 (black and gray lines, respectively) with single-integrator dynamics and controllers (5.28) (solid lines) and unicycle dynamics with controllers (5.31) (dashed lines). Square markers indicate the initial positions of the two agents, whereas circular markers indicate the points along the trajectories where the distance between the agents is at a minimum. Arrows indicate the direction of travel.

In what follows the solution to *one* optimal control problem is considered to illustrate the result of Propositions 5 and 6, before a *sequence* of such problems is considered to demonstrate how the problem of optimal monitoring is solved.

5.8.1 Approximate Solution to One Optimal Control Problem

First, consider the solution of *one* optimal control problem defined by (5.11) and (5.13) for $q_1(0) = [-8, -8]^\top$, $q_2(0) = [8, 8]^\top$, $\tilde{q}_1 = q_2(0)$ and $\tilde{q}_2 = q_1(0)$: the virtual players are such that the agents interchange positions. The situation in which both agents have unicycle dynamics (5.30) is then considered for the same initial positions and $\phi_1(0) = \frac{\pi}{8}$ and $\phi_2(0) = -\pi$.

The solid lines in Figure 5.5 shows the trajectories of agents 1 and 2 (black and gray lines, respectively), assuming both have dynamics (5.5). The solid squares indicate the initial positions of the two agents. The dashed lines show the trajectories of agents 1 and 2 (black and gray, respectively) assuming they have the unicycle dynamics (5.30). Along the trajectories, the points where the distance between the agents are at a minimum are

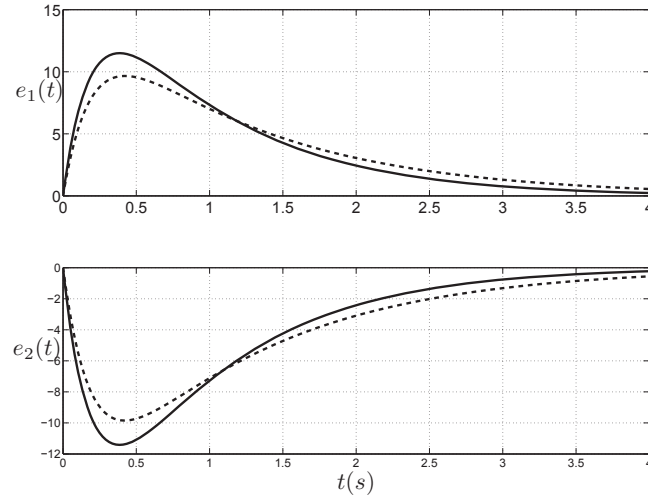


Figure 5.6: Time histories of the errors $e_1(t)$ (top) and $e_2(t)$ (bottom). The solid lines represent the x -components and the dashed lines the y -components.

identified by the solid circles, the circles centred at these points are of radius r and R , $i = 1, 2$, and show that the agents do not collide. The arrows indicate the direction of travel.

Finally, the time histories of the distance between the two agents with single-integrator dynamics (solid line) and unicycle dynamics (dashed line) are shown in Figure 5.7. The dash-dotted lines indicate the values of the risk distance, R (black line), and the safety distance, r (gray line).

5.8.2 Monitoring by Solving a Sequence of Optimal Control Problems

Consider now the case in which $q_1(0) = [-8, -8]^\top$, $q_2(0) = [8, 8]^\top$, $\tilde{q}_1 = [6, 0]^\top$, $\tilde{q}_2 = [-6, 0]^\top$ and the initial headings of the agents with the unicycle dynamics are $\phi_1(0) = -\frac{\pi}{4}$ and $\phi_2(0) = \frac{\pi}{2}$. A sequence of fifty optimal control problems have been solved to approximate the differential game (5.7). Between successive optimal control problems the values taken by the virtual players are selected according to (5.16), with $R_{d,1} = R_{d,2} = 2$.

Figure 5.8 shows the trajectories of agents 1 and 2 (black and gray lines, respectively) with dynamics (5.5) and controllers (5.28), $i = 1, 2$, (solid lines) and with dynamics

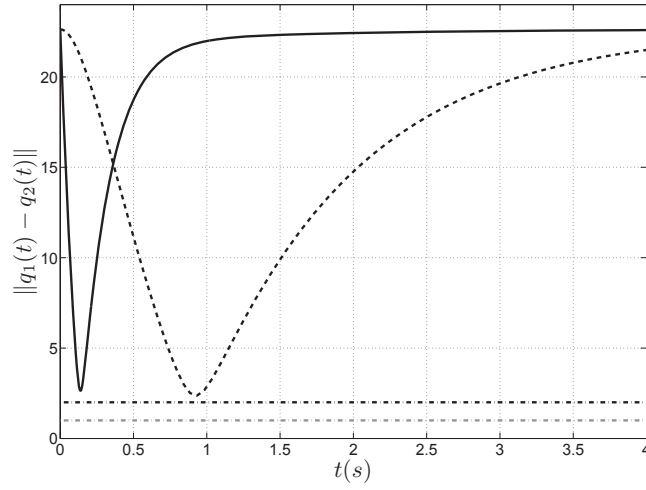


Figure 5.7: Time history of the distance between the two agents with single-integrator dynamics (solid line) and unicycle-dynamics (dashed line).

(5.30) and the controllers (5.31), $i = 1, 2$, (dashed lines). The square markers indicate the initial positions of the agents. It is clear that the agents with unicycle dynamics attempt to track the *trajectory plan*.

Figure 5.9 shows the time histories of the x - and y -components (solid and dashed lines, respectively) of the errors $e_1(t)$ (top) and $e_2(t)$ (bottom). As can be seen, errors occur when *infeasible trajectories* are generated by the *trajectory plan*, which are more likely to happen at each update of the virtual players. However, during each optimal control problem the errors converge to zero in accordance with Proposition 6.

Figure 5.10 shows the coverage map (5.1) at the end of the final optimal control problem for the case in which the agents are assumed to have unicycle dynamics. Finally, Figure 5.11 shows time histories of the maximum and minimum coverage levels (solid and dashed lines, respectively) across the region Ω . Both levels are monotonically increasing, which implies that the region is being monitored continuously.

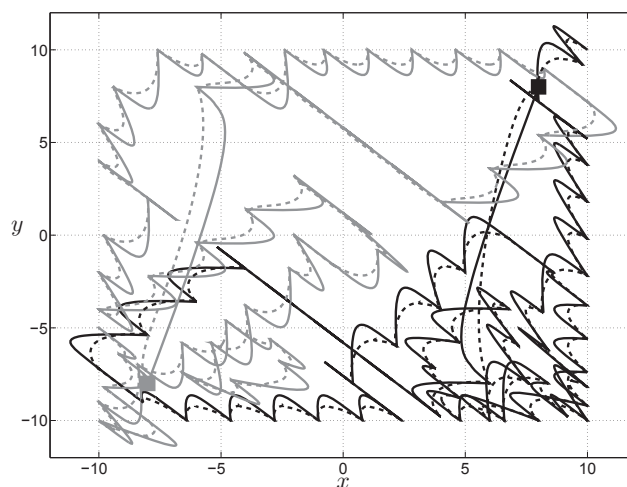


Figure 5.8: Trajectories of agents 1 and 2 (black and gray lines, respectively) with single-integrator dynamics and controllers (5.28) (solid lines) and unicycle dynamics with controllers (5.31) (dashed lines) for a sequence of fifty optimal control problems. The square markers indicate the initial positions of the two agents.

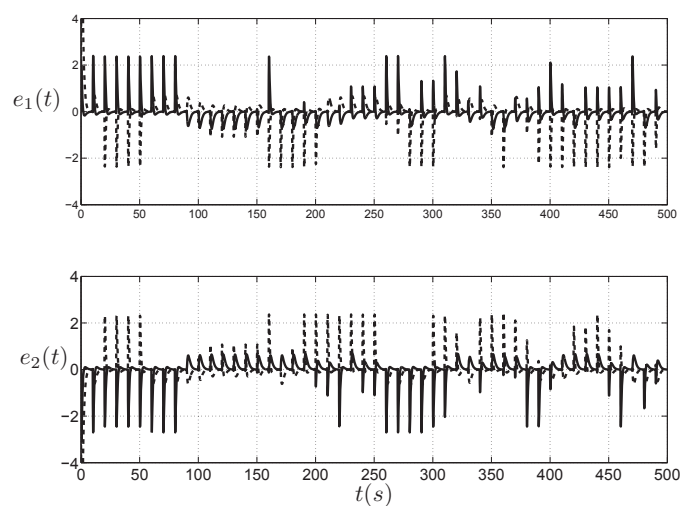


Figure 5.9: Time histories of the errors $e_1(t)$ (top) and $e_2(t)$ (bottom). The solid lines represent the x -components and the dashed lines the y -components.

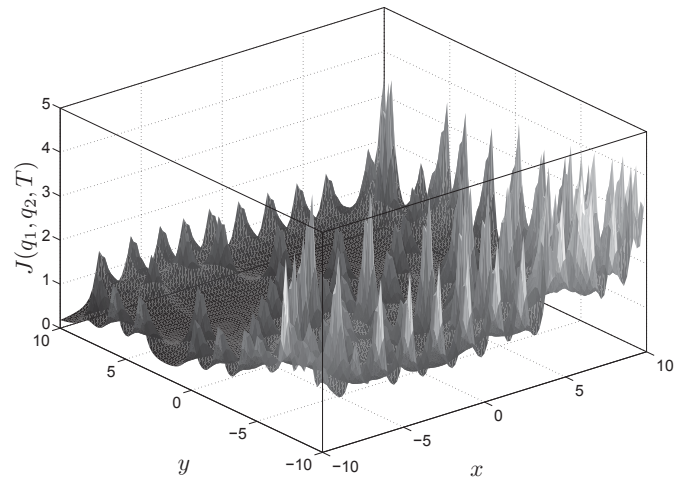


Figure 5.10: The coverage map at the end of the final optimisation problem.

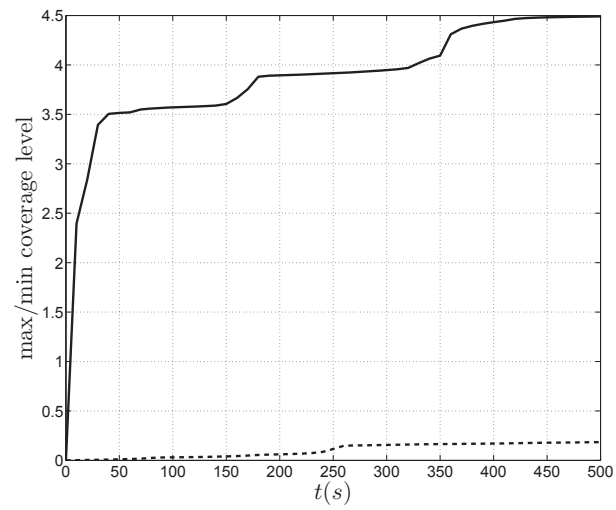


Figure 5.11: Time histories of the maximum and minimum (solid and dashed lines, respectively) coverage levels in Ω .

5.9 Conclusion

The problem of monitoring a region using a team of N sensors can be formulated as a differential game by introducing N *virtual players*. The differential game can be approximated as a sequence of infinite-horizon optimal control problems. Assuming the agents have single-integrator dynamics, closed-form solutions of the optimal control problems can be found when $N = 1$. However, this is not the case for $N > 1$ and approximate solutions must be sought. This can be done using a dynamic controller and the notion of *algebraic $\bar{P}$ matrix solution* as shown in Section 5.6. One way of extending the results to situations in which the agents have more general dynamics is to use the results based-on the single-integrator dynamics as a *trajectory plan*. For the case in which the agents have unicycle dynamics controllers that track the *trajectory plan* have been designed. Simulations have been presented for the case in which $N = 2$. First the solution of a single optimal control problem has been considered for the cases in which the agents have single-integrator dynamics and unicycle dynamics. Finally a sequence of fifty optimal control problems, which approximate the differential game, has been considered.

Chapter 6

Mean-Field Games

6.1 Introduction

It has already been discussed how differential games introduce the notion of strategic behaviour to control theory and how this can be useful for problems involving multi-agent systems. With applications in economics and power systems, to mention but a few, it is not difficult to think of problems which involve a large number of players. Recall, however, that the solution of a N -player nonzero-sum differential game relies on the solution of N coupled PDEs. Thus, as N becomes large obtaining solutions to the differential games becomes increasingly cumbersome, even when settling for approximate solutions, such as those introduced in Chapters 3 and 4. When considering differential games in which N is large we start moving into the realm of *mean-field games*. Albeit being a relatively new field of research, this branch of game theory is of practical interest in the study of a variety of problems [28, 108].

Living at the interface of mathematical physics and differential games, mean-field games deals with the study of differential games with infinitely many *indistinguishable* players [108]. As N grows, it becomes increasingly meaningful to study the behaviour of the *population* of players, often referred to as agents, instead of considering each player individually, and this *collective behaviour* is captured by a density distribution function. Since the agents are indistinguishable, they are homogeneous and have the same objec-

tives. The *mean-field approach* uses ideas from physics, more precisely from *mean-field theory*, where complex models consisting of a large number of *small*, interacting components are described by an *average effect*, which is referred to as the *mean-field*. The same idea is applied to the differential game setting: each of the agents attempts to minimise the same cost functional, and both this cost functional and the state dynamics may depend on the *average behaviour* of the agents [109]. Thus, in place of N coupled PDEs governing the differential game the problem can be equivalently described by two coupled PDEs, namely a single Hamilton-Jacobi-Bellman equation and a Fokker-Planck-Kolmogorov equation. The former is similar to the equations encountered in optimal control, whereas the latter describes the evolution of the density distribution of the agents. Thus, when considering certain types of differential games, namely those in which there are a large number of homogeneous players the framework provided by mean-field games may be useful. In particular the solution to N -player differential game problems can be approximated by a mean-field game. Furthermore, as will become apparent in Section 6.2, the approximation error scales as a function of N and the approximation becomes exact as N tends to infinity. Although the problem is reduced to solving two intertwined PDEs, an additional complication arises in this setting, namely the Hamilton-Jacobi-Bellman PDE runs backwards in time whereas the Fokker-Planck-Kolmogorov PDE runs forwards in time. This forward-backward structure is typical of mean-field games.

The remainder of the chapter is structured as follows. A brief background on mean-field games is provided in Section 6.2. Some results regarding solutions for certain mean-field games are presented in Section 6.3 before the theory is illustrated by numerical examples provided in Section 6.4. Finally some concluding remarks and directions for future work are given in Section 6.5.

The standard framework used to study mean-field games is adopted in this Chapter. When considering mean-field games it is common to include stochastic variables. Thus, some results from stochastic control theory are recalled in Section 6.3. However, a detailed background is not included in the thesis. For details on stochastic control theory see, for example, [34].

We denote with $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space. Let $\mathcal{B}$ be a finite-dimensional Brownian motion defined on this probability space. Let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be its natural filtration augmented by all the $\mathbb{P}$ -null sets (sets of measure-zero with respect to $\mathbb{P}$). We use ∂_x and ∂_{xx}^2 to denote the first and second partial derivatives with respect to x , respectively.

6.2 Background

Mean-field theory plays an important role in statistical physics and provides a method of studying large stochastic models, for example when considering systems consisting of a large number of particles. Often there are too many particles to consider the interactions between each particle individually. Thus, *mean-fields*, which describe the inter-particle interactions, are introduced and used to construct approximations of the systems behaviour. The idea behind the introduction of the mean-fields is that the effect of the population on an individual can be approximated by an *average* effect, which is described by a mean-field [108, 109]. The theory of mean-field games, which was introduced independently by J.M. Lasry and P.L. Lions in [28] and by M.Y. Huang, P.E. Caines and R.P. Malhamé in [29, 30], draws inspiration from this. Differential games with a large number of homogeneous players which interact *strategically* are considered and using tools from mean-field theory the effect all the players have on an individual is approximated by an *average* effect, which is dependent on the density distribution of the players. Thus, in place of N coupled PDEs, the differential game is described by two PDEs, namely the Hamilton-Jacobi-Bellman PDE, which relates to each individual agent, and the Fokker-Planck-Kolmogorov PDE which describes the evolution of the density distribution of the agents. Each individual player then constructs its strategy based on its own state and information relating to the distribution of the rest of the population.

Consider a differential game with infinitely many players, also referred to as agents, with a scalar state $x(t) \in \mathbb{R}$. The density distribution of the agents at time t is

given according to a scalar function $m_t(x)$, namely

$$\begin{aligned} m : \mathbb{R} \times [0, +\infty[&\rightarrow [0, +\infty[, (x, t) \mapsto m_t(x) \\ m_t(x) &\geq 0 \\ \int_{\mathbb{R}} m_t(x) dx &= 1 \text{ for every } t. \end{aligned}$$

Furthermore, suppose the initial density distribution of the agents is given by $m_0(x)$. The density distribution gives a macroscopic description of the game. The states of each agent evolve according to the dynamics

$$dx_t = [f(x) + g(x)u] dt + \sigma [x dB + \zeta dt], \quad (6.1)$$

where $u(t) \in \mathbb{R}$ is the control strategy of an agent, $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ and $g(x) : \mathbb{R} \rightarrow \mathbb{R}$ are mappings, $\zeta(t) \in \mathbb{R}$ is an adversarial disturbance, $B(t)$ is a Brownian motion, which is independent of the initial state $x(0) = x_0$, and $\sigma \in \mathbb{R}$. Over a time period $[0, T]$, the agents seek to minimise the cost functional

$$J(x, m(x), u) = \int_0^T c(x, u, m) dt - \gamma^2 \int_0^T \|\zeta\|^2 dt + g(x_T, m_T), \quad (6.2)$$

where $x_T = x(T)$, $c(x, m(x), u)$ is a running cost, which may depend on both the state of the agent and the overall distribution of the agents, the second term is a quadratic penalty on the unknown disturbance and the final term is a terminal penalty, which, again, may depend on both the state of the agent and the distribution $m_t(x)$. The mean-field game is then defined as follows.

Problem 12. Consider the system of agents with dynamics (6.1) and cost functionals (6.2). Suppose that the initial conditions, x_0 , of each agent are such that their density distribution is given by $m_0(x)$. Let $m_t^*(x)$ be the optimal mean-field trajectory. The mean-field game lies in determining the strategies, $u(t)$, solving

$$\inf_u \sup_{\zeta} J(x, u, m^*, \zeta)$$

subject to the dynamics (6.1).

Problem 12 is known as a *robust mean-field game*. Note that Problem 12 describes a zero-sum game between the agent and the disturbance. This is similar to what is often seen in H_∞ control.

Let $H(x, \lambda, m_t)$ denote the Hamiltonian, namely

$$H(x, u, \lambda, m_t, \zeta) = c(x, u, m) - \gamma^2 \zeta^2 + \lambda(f(x) + g(x)u + \sigma(\zeta(t))),$$

where $\lambda \in \mathbb{R}$ is the costate.

Theorem 9. Suppose we can find a value function $V_t(x)$ and a density distribution $m_t(x)$ satisfying

$$\begin{aligned} \dot{V}_t + H\left(x, \frac{\partial V_t}{\partial x}, u^*, m_t, \zeta^*\right) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 V_t}{\partial x^2} &= 0, \text{ in } \mathbb{R} \times [0, T[, \\ V_T(x) &= g(x, m), \text{ in } \mathbb{R}, \\ \frac{\partial m_t}{\partial t} + \operatorname{div}_x \left(m_t \frac{\partial H}{\partial \lambda} \Big|_{\lambda = \frac{\partial V}{\partial x}} \right) - \frac{1}{2} \sigma^2 \frac{\partial^2 x^2 m_t}{\partial x^2} &= 0, \text{ in } \mathbb{R} \times [0, T[, \\ m_0(x), \quad &\text{given.} \end{aligned} \tag{6.3}$$

Then the worst-case disturbance is given by

$$\zeta^* = \frac{\sigma}{2\gamma^2} \frac{\partial V}{\partial x},$$

and the optimal strategies of each agent are given by

$$u^* = \inf_u H\left(x, u, \frac{\partial V}{\partial x}, m_t, \zeta^*\right).$$

◇

The first equation in (6.3) is the Hamilton-Jacobi-Bellman PDE, whereas the second equation is the Fokker-Planck-Kolmogorov PDE. For more detail and the proof see, for example, [28, 108, 109]. The mean-field game PDEs have a particular structure: the Hamilton-Jacobi-Bellman is solved backwards, as is the case with optimal control problems and *standard* N -player Nash differential games. Thus $m_t(x)$ must be given to solve the

Hamilton-Jacobi-Bellman PDE. On the other hand, the second equation, namely the Fokker-Planck-Kolmogorov PDE, is solved forwards in time, with $m_0(t)$ specified at the beginning of the problem. This forward-backward structure is typical of mean-field games [108].

Solutions to the two coupled PDEs in (6.3), *i.e.* V_t and m_t with the given boundary conditions on V_T and m_0 must be obtained, which is not generally straight-forward. This is commonly done by iteratively solving the Hamilton-Jacobi-Bellman equation for fixed m_t and by entering the corresponding optimal control strategy u^* in the Fokker-Planck-Kolmogorov equation in (6.3), until a fixed point in V_t and m_t is reached. Any solution of the system of equations (6.3) is referred to as *worst-disturbance feedback mean-field equilibrium*.

In some cases solutions to the PDEs can be obtained. Notably, a class of linear-quadratic mean field games and their solutions are discussed in [109, 110]. Furthermore, it is shown in the following section that in some scenarios the PDEs (6.3) can be replaced by a system of ODEs.

6.3 Mean-Field Games and Two-Point Boundary Value Problems

In this section a class of mean-field games is considered. In line with the theory of mean-field games, a population of indistinguishable dynamic agents, the dynamics of which are given by a linear stochastic differential equation driven by a Brownian motion and under the influence of a control and an adversarial disturbance, is considered. We consider the problem in which the population consists of so-called “crowd-averse” dynamic agents, *i.e.* the agents are such that they seek to regulate their state to values characterised by a low density, thus avoiding “crowded” states. Similar problems arise naturally in social sciences: the states represent opinions, the dynamics represent the propagation of these opinions, and crowd-averse attitudes capture the agents’ willingness to escape consensus and seek dissensus. The crowd-averse behaviour is described by a cost functional which

involves quadratic penalty on control and mean-field term involving the density of the players. This problem, which includes both the *individual* agent and the behaviour of the *population* can be classified as a mean-field game, similar to Problem 12. The case in which the initial distribution is a sum of polynomial terms and the value function is quadratic is analysed. Problems similar to the one studied in the following sections arise in different fields, for example in opinion dynamics in social networks [111]. Crowd-averse attitudes in this setting imply that the players tend to have very different opinions. This is in contrast to the “opposite” phenomena of “emulation”, “mimicry” or “herd behavior”. The consideration of crowd-averse dynamic agents may also be of interest in problems regarding crowd dynamics and pedestrian flows [112].

6.3.1 Problem Formulation

Consider a game with infinitely many homogeneous players, also referred to as agents. Since the players are homogeneous they are indistinguishable. For each player let $x(0) = x_0$ be its initial state, which is realised according to the probability distribution $m_0(x)$. The state of the player at time t denoted by $x(t) \in \mathbb{R}$ evolves according to a controlled stochastic process over a finite horizon $T > 0$, i.e. it satisfies the dynamics

$$dx = [\alpha x + \beta u] dt + \sigma [x d\mathcal{B} + \zeta dt] \quad (6.4)$$

where $u(t) \in \mathbb{R}$ is the control input, $\mathcal{B}(t) \in \mathbb{R}$ is a Brownian motion, which is independent of the initial state x_0 and independent across players and time. The constants $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$ and $\sigma \in \mathbb{R}$ are parameters, and $\zeta_t \in \mathbb{R}$ is an adversarial disturbance. This is a class of the more general dynamics (6.1) in which the state is scalar and satisfies linear dynamics.

Consider the following assumption on the distribution at time t , $m_t(x)$.

Assumption 13. The density distribution $m_t(x)$ has compact support and within its sup-

port it is given by

$$\begin{aligned} m_t(x) &= a_{0t} + \sum_{j=1}^m \frac{1}{j} a_{jt} x^j, \text{ in } \mathbb{R} \times [0, T] \\ m_0(x) &= a_{00} + \sum_{j=1}^m \frac{1}{j} a_{j0} x^j, \\ a_{j0} &\text{ given for all } j = 0, \dots, m. \end{aligned} \tag{6.5}$$

Remark 30. The sum of polynomial terms in Assumption 13 can be interpreted as the m^{th} degree Taylor approximation of a general distribution $m_t(x)$. ▲

Each agent is associated with a cost functional of the form (6.2). We consider the running cost

$$c(x, u, m) = a_{0t} + a_{1t}x + \frac{1}{2}a_{2t}x^2 + \frac{b}{2}u^2. \tag{6.6}$$

The first three terms represent the mean-field cost and their sum is the second order Taylor approximation of the density distribution, whereas the last term, with $b > 0$, accounts for a penalty on the control energy. The penalty on the final state is

$$g(x_T, m_T(x_T)) = a_{0T} + a_{1T}x + \frac{1}{2}a_{2T}x^2, \tag{6.7}$$

namely it is a penalty on the second order Taylor approximation of the state density distribution at the end of the horizon.

Remark 31. The cost function (6.2) with running cost (6.6) and terminal penalty (6.7) is such that each of the agents are crowd-averse in the sense that they seek to regulate their states to regions of low density. Note, however, that the agents only consider the density in the state they are in: this is known as *local interaction*. Thus, it may happen that all agents attempt to regulate their state to the same low-density state. ▲

We then consider the following problem.

Problem 13. Consider the system of agents with linear dynamics (6.4) and cost functionals (6.2) with running cost (6.6) and terminal penalty (6.7). Suppose the initial conditions, x_0 , of each agent are such that their density distribution is given by $m_0(x)$. Let $m_t^*(x)$ be the optimal mean-field trajectory. Then, solving the mean-field game consists in deter-

mining the feedback strategies u solving

$$\inf_u \sup_{\zeta} J(x, u, m^*, \zeta),$$

subject to the dynamics (6.4).

As mentioned in Section 6.2, any solution of the system of equations (6.3) is referred to as *worst-disturbance feedback mean-field equilibrium*. The existence of solutions for problem (6.3) can be guaranteed under the following assumptions. Let the initial probability distribution m_0 be absolutely continuous with a finite second moment. As the integrand of the cost is convex in u , and concave in the disturbance ζ , one gets a convex-concave stage cost function. The drift is linear and hence Lipschitz continuous because α, β, σ are bounded. We assume that the Fenchel transform of c is Lipschitz in (x, z) . Finally, we assume that the function $p \mapsto \frac{\sigma^2}{4\gamma^2} \|p\|^2 + H$ is strictly convex, differentiable and $\frac{\sigma^2}{4\gamma^2} \|p\|^2 + H$ is Lipschitz continuous. Note that this last condition is weaker than the convexity assumption on H . Under the above main assumptions, the existence of a solution is established in Theorem 2.6 in [28]. In addition to this, as the cost is Lipschitz continuous in m_t the solution to the asymptotic case with infinitely many players is related to the case with a finite number of players N by the classical bound $\frac{1}{N}$ provided in [28, 30, 108]. Thus, the solution to a mean-field game can be used to approximate a similar differential game with a finite number of players and the approximation improves as the number of players increases.

6.3.2 Main results

Let $V_t(x)$ be the (upper) value of the robust optimization problem under worst-case disturbance starting at time t from state x . We consider quadratic value functions of the form

$$\begin{aligned} V_t(x) &= q_{0t} + \sum_{j=1}^2 \frac{1}{j} q_{jt} x^j, \text{ in } \mathbb{R} \times [0, T] \\ V_T(x) &= g(x_T, m_T(x_T)) = q_{0T} + \sum_{j=1}^2 \frac{1}{j} q_{jT} x^j = a_{0T} + \sum_{j=1}^2 \frac{1}{j} a_{jT} x^j. \end{aligned} \quad (6.8)$$

The case of a crowd-averse system in which the players seek to drive their state towards state values characterized by a lower density, namely Problem 13 is considered.

Theorem 10. The mean-field system associated to the robust mean-field game for the crowd-averse system is described by the equations:

$$\begin{aligned}
& \dot{V}_t + \left(-\frac{\beta^2}{2b} + \left(\frac{\sigma}{2\gamma} \right)^2 \right) \left(\frac{\partial V_t}{\partial x} \right)^2 + \alpha x \frac{\partial V_t}{\partial x} + a_{0t} + a_{1t}x + \frac{1}{2}a_{2t}x^2 \\
& + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V_t}{\partial x^2} = 0, \text{ in } \mathbb{R} \times [0, T[, \\
& V_T(x) = q_{0T} + \sum_{j=1}^2 \frac{1}{j} q_{jT} x^j, \text{ in } \mathbb{R}, \\
& \frac{\partial m_t}{\partial t} + \sum_{j=1}^n a_{jt} \left(\left(1 + \frac{1}{j} \right) \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) x_t^j + \left(-\frac{\beta^2}{b} q_{1t} + \frac{\sigma^2}{2\gamma^2} q_{1t} \right) x_t^{j-1} \right) \\
& + a_{0t} \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) - \frac{1}{2}\sigma^2 \frac{\partial^2 x^2 m_t}{\partial x^2} = 0, \text{ in } \mathbb{R} \times [0, T[, \\
& m_0(x) = a_{00} + \sum_{j=1}^n \frac{1}{j} a_{j0} x^j \text{ in } \mathbb{R}.
\end{aligned} \tag{6.9}$$

The optimal control and worst disturbance are then given by

$$\begin{aligned}
u_t^* &= -\frac{\beta}{b} \frac{\partial V_t}{\partial x}, \\
\zeta_t^* &= \frac{\sigma}{2\gamma^2} \frac{\partial V_t}{\partial x}.
\end{aligned} \tag{6.10}$$

◇

Proof: The proof consists of two parts. The first part consists in showing that condition (6.10) holds. It is then shown that (6.9) holds.

The Hamiltonian associated with Problem 13 is of the form

$$H\left(x_t, \frac{\partial V_t}{\partial x}, m_t\right) = a_{0t} + a_{1t}x + \frac{1}{2}a_{2t}x^2 + \frac{b}{2}u^2 - \gamma^2\zeta^2 + \frac{\partial V_t}{\partial x}(\alpha x_t + \beta u + \sigma\zeta). \tag{6.11}$$

Differentiating with respect to u and ζ under the assumption of concavity on ζ gives the

conditions for the optimal control and the worst case disturbance, respectively, *i.e.*

$$bu^* + \frac{\partial V_t}{\partial x} \beta = 0, \quad (6.12)$$

$$-2\gamma^2 \zeta^* + \frac{\partial V_t}{\partial x} \sigma = 0. \quad (6.13)$$

Thus, the first part of the proof is concluded by noting that (6.12) and (6.13) imply (6.10).

We now turn to the second part of the proof, namely showing that (6.9) holds. First, note that the mean-field system associated with the robust mean-field game introduced in Problem 13 is given by

$$\begin{aligned} \frac{\partial V}{\partial t} + H\left(x, \frac{\partial V_t}{\partial x}, m_t\right) + \left(\frac{\sigma}{2\gamma}\right)^2 \left(\frac{\partial V_t}{\partial x}\right)^2 + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V_t}{\partial x^2} &= 0, \text{ in } \mathbb{R} \times [0, T[, \\ V_T(x) &= g(x, m), \text{ in } \mathbb{R}, \\ \frac{\partial m_t}{\partial t} + \frac{\partial}{\partial x} \left(m_t \frac{\partial H(x, \lambda, m)}{\partial \lambda} \Big|_{\lambda = \frac{\partial V_t}{\partial x}} \right) - \frac{1}{2}\sigma^2 \frac{\partial^2 x^2 m_t}{\partial x^2} &= 0, \text{ in } \mathbb{R} \times [0, T[, \\ m_0(x) &\text{ given,} \end{aligned} \quad (6.14)$$

which follows directly from (6.3) since the state x is scalar.

The second and last equations of (6.9) are the boundary conditions and derive straightforwardly from Bellman equations and the evolution of the state. The first equation of (6.9) follows from substituting (6.10) in the Hamiltonian (6.11), yielding

$$H\left(x_t, \frac{\partial V_t}{\partial x}, m_t\right) = a_{0t} + a_{1t}x + \frac{1}{2}a_{2t}x^2 - \frac{\beta^2}{2b} \left(\frac{\partial V_t}{\partial x}\right)^2 + \alpha x \frac{\partial V_t}{\partial x} + \frac{\sigma^2}{4\gamma^2} \left(\frac{\partial V_t}{\partial x}\right)^2.$$

Substituting the above expression into (6.14), the Hamilton-Jacobi-Bellman equation in (6.9) is obtained.

Finally, the third equation of (6.9), which is a PDE representing the Fokker-Planck-Kolmogorov equation, is proved by substituting (6.10) into the Fokker-Planck-Kolmogorov equation in (6.14) and expanding the divergence operation, namely

$\frac{\partial(mdx)}{\partial x} = \frac{\partial m}{\partial x}dx + m\frac{\partial(dx)}{\partial x}$. This yields

$$\frac{\partial m_t}{\partial t} + \left(\alpha x - \frac{\beta^2}{b} \frac{\partial V_t}{\partial x} + \frac{\sigma^2}{2\gamma^2} \frac{\partial V_t}{\partial x} \right) \frac{\partial m_t}{\partial x} + m_t \left(\alpha - \frac{\beta^2}{b} \frac{\partial^2 V_t}{\partial x^2} + \frac{\sigma^2}{2\gamma^2} \frac{\partial^2 V_t}{\partial x^2} \right) - \frac{1}{2} \sigma^2 \frac{\partial^2 x^2 m_t}{\partial x^2} = 0.$$

Using Assumption 13 and (6.8) it follows that

$$\frac{\partial m_t}{\partial x} = \sum_{j=1}^n a_{jt} x^{j-1},$$

and

$$\frac{\partial V_t}{\partial x} = \sum_{j=1}^2 q_{jt} x^{j-1} = q_{2t}x + q_{1t}.$$

By substitution

$$\begin{aligned} \left(\alpha x - \frac{\beta^2}{b} \frac{\partial V_t}{\partial x} + \frac{\sigma^2}{2\gamma^2} \frac{\partial V_t}{\partial x} \right) \frac{\partial m_t}{\partial x} &= \sum_{j=1}^n a_{jt} x^{j-1} \left(\alpha x - \frac{\beta^2}{b} (q_{2t}x + q_{1t}) + \frac{\sigma^2}{2\gamma^2} (q_{2t}x + q_{1t}) \right) \\ &= \sum_{j=1}^n a_{jt} \left(\left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) x^j + \left(-\frac{\beta^2}{b} q_{1t} + \frac{\sigma^2}{2\gamma^2} q_{1t} \right) x^{j-1} \right), \end{aligned}$$

and

$$\begin{aligned} m_t \left(\alpha - \frac{\beta^2}{b} \frac{\partial^2 V_t}{\partial x^2} + \frac{\sigma^2}{2\gamma^2} \frac{\partial^2 V_t}{\partial x^2} \right) &= a_{0t} \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) \\ &\quad + \sum_{j=1}^n \frac{1}{j} a_{jt} \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) x^j, \end{aligned}$$

and the Fokker-Planck-Kolmogorov equation in (6.9) follows, which concludes the proof. $\square$

In the following statement it is established that the mean-field system (6.9) can be replaced by a two-point boundary value problem.

Theorem 11. The mean-field system associated to the robust mean-field game for the crowd-averse system is equivalently described by the ordinary differential equations:

$$\begin{aligned}
\dot{q}_{0t} + \left(-\frac{\beta^2}{2b} + \left(\frac{\sigma}{2\gamma} \right)^2 \right) q_{1t}^2 + a_{0t} &= 0, \\
\dot{q}_{1t} + \left(-\frac{\beta^2}{2b} + \left(\frac{\sigma}{2\gamma} \right)^2 \right) 2q_{1t}q_{2t} + \alpha q_{1t} + a_{1t} &= 0, \\
\frac{1}{2}\dot{q}_{2t} + \left(-\frac{\beta^2}{2b} + \left(\frac{\sigma}{2\gamma} \right)^2 \right) q_{2t}^2 + \alpha q_{2t} + \frac{1}{2}a_{2t} + \frac{\sigma^2}{2}q_{2t} &= 0, \\
q_{jT} &= a_{jT}, \\
\dot{a}_{0t} + a_{1t} \left(-\frac{\beta^2}{b}q_{1t} + \frac{\sigma^2}{2\gamma^2}q_{1t} \right) + a_{0t} \left(\alpha - \frac{\beta^2}{b}q_{2t} + \frac{\sigma^2}{2\gamma^2}q_{2t} \right) - \frac{1}{2}\sigma^2 2a_{0t} &= 0, \\
\dot{a}_{1t} + a_{1t} \left(\alpha 2 + \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) 2q_{2t} \right) + a_{2t} \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) q_{1t} - \frac{6}{2}\sigma^2 a_{1t} &= 0, \\
\frac{1}{j}\dot{a}_{jt} + a_{jt} \left(\alpha \left(1 + \frac{1}{j} \right) + \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) \left(1 + \frac{1}{j} \right) q_{2t} \right) \\
+ a_{j+1t} \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) q_{1t} - \frac{1}{2}\sigma^2 \frac{(j+2)(j+1)}{j} a_{jt} &= 0, \quad j = 2, \dots, n-1 \\
\frac{1}{n}\dot{a}_{nt} + a_{nt} \left(1 + \frac{1}{n} \right) \left(\alpha - \frac{\beta^2}{b}q_{2t} + \frac{\sigma^2}{2\gamma^2}q_{2t} \right) - \frac{1}{2}\sigma^2 \frac{(n+2)(n+1)}{n} a_{nt} &= 0 \\
a_{j0} \text{ given for all } j = 1, \dots, n.
\end{aligned} \tag{6.15}$$

The optimal control and worst disturbance are then given by

$$\begin{aligned}
u^* &= -\frac{\beta}{b}(q_{2t}x_t + q_{1t}), \\
w^* &= \frac{\sigma}{2\gamma^2}(q_{2t}x_t + q_{1t}).
\end{aligned} \tag{6.16}$$

◇

Proof: It follows from Assumption 13 and the expression for the value function in (6.8) that

$$\frac{\partial V_t}{\partial t} = \dot{q}_{0t} + \dot{q}_{1t}x + \frac{1}{2}\dot{q}_{2t}x^2,$$

and

$$\frac{\partial m_t}{\partial t} = \dot{a}_{0t} + \sum_{j=1}^n \frac{1}{j} \dot{a}_{jt} x^j.$$

Furthermore,

$$\frac{\partial^2 x^2 m_t}{\partial x^2} = 2a_{0t} + \sum_{j=1}^n \frac{(j+2)(j+1)}{j} a_{jt} x^j.$$

Substituting these expressions into (6.9) yields

$$\begin{aligned} & \dot{q}_{0t} + \sum_{j=1}^2 \frac{1}{j} \dot{q}_{jt} x^j + \left(-\frac{\beta^2}{2b} + \left(\frac{\sigma}{2\gamma} \right)^2 \right) (q_{1t} + q_{2t} x_t)^2 + \sum_{j=1}^2 \alpha q_{jt} x^j + a_{0t} \\ & + \sum_{j=1}^n \frac{1}{j} a_{jt} x^j + \frac{\sigma^2}{2} q_{2t} x^2 = 0, \text{ in } \mathbb{R} \times [0, T[, \\ & v_T(x) = q_{0T} + \sum_{j=1}^2 \frac{1}{j} q_{jT} x^j, \text{ in } \mathbb{R}, \\ & \dot{a}_{0t} + \sum_{j=1}^n \frac{1}{j} \dot{a}_{jt} x^j + a_{jn} \left(1 + \frac{1}{n} \right) \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) x_t^n \\ & + \sum_{j=1}^{n-1} a_{jt} \left(\alpha \left(1 + \frac{1}{j} \right) + \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) \left(1 + \frac{1}{j} \right) q_{2t} \right) x_t^j \\ & + \sum_{j=1}^{n-1} a_{jt} a_{j+1t} \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) q_{1t} x_t^j + a_{1t} \left(-\frac{\beta^2}{b} q_{1t} + \frac{\sigma^2}{2\gamma^2} q_{1t} \right) \\ & + a_{0t} \left(\alpha - \frac{\beta^2}{b} q_{2t} + \frac{\sigma^2}{2\gamma^2} q_{2t} \right) - \frac{1}{2} \sigma^2 \left(2a_{0t} + \sum_{j=1}^n \frac{(j+2)(j+1)}{j} a_{jt} x^j \right) \\ & = 0, \text{ in } \mathbb{R} \times [0, T[, \\ & m_0(x) = a_{00} + \sum_{j=1}^n \frac{1}{j} a_{j0} x^j \text{ in } \mathbb{R}. \end{aligned} \tag{6.17}$$

As the above equations must hold for all $x \in \mathbb{R}$, the coefficients of each power of x must cancel. Thus, collecting the terms of the same powers of x , the ODEs in (6.15) are obtained.

Finally, note that $V_t = q_{0t} + q_{1t}x + \frac{1}{2}q_{2t}x^2$ is the solution to the Hamilton-Jacobi-Bellman equation in (6.9). The optimal control and worst-case disturbance (6.16) then follows from (6.10) and this concludes the proof. $\square$

Remark 32. The ODEs in (6.15) constitute a somewhat atypical two-point boundary value problem since the initial conditions a_{j0} are given for all j , whereas the final conditions $q_{jT} = a_{jT}$ for $j = 0, 1, 2$ are unknown a-priori. However, by performing the change of coordinates $\tilde{q}_{jt} = q_{jt} - a_{jt}$ for $j = 0, 1, 2$, the problem can be transformed into a standard two-point boundary value problem with final conditions $\tilde{q}_{jT} = 0$. Solutions to this modified problem can be found numerically, for example using the shooting method, and these can be used to obtain solutions to the original ODEs (6.15). $\blacktriangle$

6.3.3 Interpretation of Results

In this section it is shown that the stochastic differential equation describing the closed-loop system has an exponentially and asymptotically stable equilibrium. Substituting the optimal control and the worst-case disturbance (6.16) into the dynamics for x yields the closed-loop system

$$dx = \left[\alpha + \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) q_t \right] x dt + \sigma x dB, \quad t \in (0, T], \quad x_0 \in \mathbb{R}.$$

In what follows it is assumed that the following property holds true.

Assumption 14. There exists $\kappa > 0$ such that

$$-\kappa x(t) \geq \left(\alpha + \left(-\frac{\beta^2}{b} + \frac{\sigma^2}{2\gamma^2} \right) q_t \right) x(t), \quad (6.18)$$

for all $t \geq 0$.

Assumption 14 ensures that the evolution of the state is bounded from above by an exponential decay. With this assumption the stability analysis can be performed within the framework of stochastic stability theory [113]. Consider the infinitesimal generator

$$\mathcal{L} = \frac{1}{2} \sigma^2 x^2 \frac{d^2}{dx^2} - \kappa x \frac{d}{dx}, \quad (6.19)$$

and the Lyapunov function $V(x) = x^2$. The stochastic derivative of $V(x)$ is obtained by

applying the infinitesimal generator to $V(x)$. This yields

$$\begin{aligned}\mathcal{L}V(x(t)) &= \lim_{dt \rightarrow 0} \frac{\mathbb{E}V(x(t+dt)) - V(x(t))}{dt} \\ &= [\sigma^2 - 2\kappa]x(t)^2.\end{aligned}$$

Proposition 7. [113] Suppose Assumption 14 holds. If $V(x) \geq 0$, $V(0) = 0$ and $\mathcal{L}V(x) \leq -\eta V(x)$ on $Q_\epsilon := \{x : V(x) \leq \epsilon\}$, for some $\eta > 0$, and for arbitrarily large ϵ , then the origin is asymptotically stable “with probability one”, and

$$P_{x_0} \left\{ \sup_{T \leq t < +\infty} x(t)^2 \geq \lambda \right\} \leq \frac{V(x_0)e^{-\psi T}}{\lambda},$$

for some $\psi > 0$. ◇

From the above theorem we have the following result, which establishes exponential stochastic stability of the mean-field equilibrium.

Corollary 2. Let Assumption 14 hold. If $[\sigma^2 - 2\kappa] < 0$ then $\lim_{t \rightarrow \infty} x(t) = 0$ almost surely and

$$P_{x_0} \left\{ \sup_{T \leq t < +\infty} x(t)^2 \geq \lambda \right\} \leq \frac{V(x_0)e^{-\psi T}}{\lambda},$$

for some $\psi > 0$. ◇

Remark 33. For the special case in which the density distribution $m_t(x)$ is purely quadratic, some observations can be made on the evolution of the density distribution, as illustrated in Figure 6.1. In particular, Figure 6.1(a) depicts the initial density distribution $m_0(x)$ as a function of the state. The density $m_0(x) = \frac{1}{2}a_{2t}x^2$ is quadratic in x (depicted by the gray area). If the vector field is converging to zero, the density function shrinks towards zero and becomes “more convex”, as illustrated in Figure 6.1(b): this corresponds to a_{2t} increasing with t . This occurs when Assumption 14 holds true, as all players are drawn towards the origin by the linear feedback. On the other hand, if the vector field is diverging from zero, the density function is drawn apart from zero and becomes “less convex” and “more flat”, which corresponds to a_{2t} decreasing with t . This is due to a higher influence on the part of the disturbances (both the stochastic one, namely the Brownian motion, and the adversarial one ζ). This case is illustrated in Figure 6.1(c),

in which the gray area is more dispersed. Furthermore, Figure 6.1(b) also illustrates the observation made in Remark 31, namely the scenario in which all agents aim to reach the same low-density point in the state space. ▲

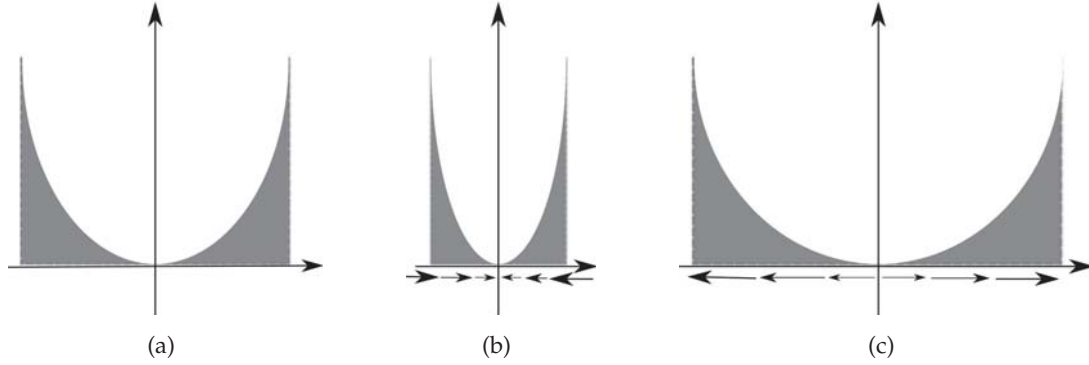


Figure 6.1: Initial distribution m_0 (a) and final distribution m_T for converging (b) and diverging vector field (c). Graphs are not to scale.

6.4 Simulations

Two numerical examples illustrating the theory are presented in this section. First, an example in which the density distribution is quadratic is then considered. This example illustrates the observations made in Remark 33. A second example, in which the distribution function is a fourth order polynomial is considered. Both examples illustrate that the agents may attempt to reach the same low-density states, as stated in Remark 31.

6.4.1 Quadratic Density Distribution

Consider a system consisting of $n = 7700$ indistinguishable players with dynamics (6.4) and suppose each of the players seek to minimise the cost functionals (6.2) with running cost (6.6) and terminal cost (6.7). Furthermore, suppose that the initial distribution is quadratic, *i.e.* $m_0(x) = \frac{1}{2}a_{20}x^2$. It follows from Theorem 10 that the optimal control and the worst-case disturbance are given by (6.16), which relies on the solution of the two-

point boundary value problem (6.15). The numerical results are obtained by solving the coupled ODEs (6.15) numerically before using the solution to simulate the closed-loop system (6.4) for a discretised set of states, namely $x \in [-1, 1]$. The states of the n agents are initially within this set. The state trajectories are computed over the period $[0, 5]$ using the sample time 0.01. The parameters used are $a_{20} = 0.2597$, $\alpha = -0.1$, $\beta = 0.1$, $b = \gamma = 1$. The simulations have been run for $\sigma = 0$, *i.e.* without noise, and $\sigma = 0.1$. Notice that in this case, since the distribution is quadratic, the value function is quadratic too.

Figure 6.2 illustrates the solution to the coupled ODEs (6.15). The solid lines show the time history of a_{2t} , whereas the dashed lines show the time history of q_{2t} , for $\sigma = 0$ (top) and $\sigma = 0.1$ (bottom). Note that the boundary conditions are satisfied, *i.e.* $a_2(0) = a_{20}$ and $q_{2T} = a_{2T}$. Furthermore, a_{2t} is monotonically increasing with time. Figures 6.3 and 6.4 show the initial (black, dashed line) and final (black, solid line) distribution of the states of the agents for $\sigma = 0$ and $\sigma = 0.1$, respectively. The black dash-dotted lines indicate the distribution of the agents at $t = 2.5$. The gray lines indicate the distribution “predicted” by the solution to the ODEs (6.15), shown in Figure 6.2 for $t = 2.5$ (dash-dotted line) and at the final time (solid line). The distribution computed based on the evolution of the states matches well with the solution of (6.15).

The time evolution of the distribution function is a shrinking quadratic function in accordance with 6.1(b). Since the initial distribution is such that $x = 0$ is the state with the lowest density, it is expected that the players move towards this point, which is consistent with Figures 6.3 and 6.4.

6.4.2 Fourth Order Polynomial Distribution

Consider a system consisting of $n = 5422$ indistinguishable agents with dynamics (6.4). Furthermore, suppose each of the agents seeks to minimise a cost functional of the form (6.2) with running cost (6.6) and terminal cost (6.7). The initial distribution is given by $m_0 = 0.0184 + 0.0184x + \frac{1}{2}0.0373x^2 + \frac{1}{3}0.0019x^3 - \frac{1}{4}0.0100x^4$, *i.e.* the initial distribution is such that Assumption 13 is satisfied. The optimal control and the worst-case disturbance are then given by (6.16), which relies on the solution of the ODEs (6.15). These ODEs

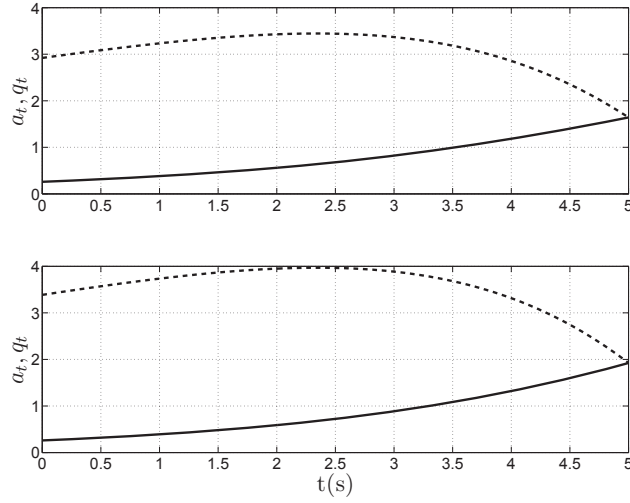


Figure 6.2: Time histories of a_{2t} (solid line) and q_{2t} (dashed line) for $\sigma = 0$ (top) and $\sigma = 0.1$ (bottom).

are solved numerically using the method discussed in Remark 32. As in the previous example, the solution of (6.15) is then used to simulate the closed-loop system (6.4) for a discretised set of states in the region $[-1, 1]$. The states of the n agents are initially within this set of states and their trajectories are computed over the period $[0, 5]$. Simulations have been run for $\sigma = 0$, *i.e.* without any noise or disturbance, and with $\sigma = 0.05$ and $\sigma = 0.1$. The remainder of the parameters are $\alpha = -0.01$, $\beta = 0.2$, $b = 0.1$ and $\gamma = 1$.

Figures 6.5 and 6.6 show the solution to the two point boundary value problem (6.9), for the case in which $\sigma = 0.05$. The solid lines in Figure 6.5 show the time history of the coefficients a_{0t} (top), a_{1t} (middle) and a_{2t} (bottom), whereas the dashed lines show the time histories of the coefficients q_{0t} (top), q_{1t} (middle) and q_{2t} (bottom). boundary conditions, $q_{jT} = a_{jT}$, are satisfied for $j = 0, 1, 2$. Figure 6.6 shows the time histories of the coefficients a_{3t} (top) and a_{4t} (bottom). The results are similar for $\sigma = 0$ and $\sigma = 0.1$.

Figures 6.7, 6.8 and 6.9 show the initial (black, dashed line) and final (black, solid line) distribution of the states of the agents for $\sigma = 0$, $\sigma = 0.05$ and $\sigma = 0.1$, respectively. The dash-dotted black lines indicate the distribution of the agents at an intermediate time, namely at $t = 2$. The gray lines indicate the distribution expected from the solution of the two-point boundary value problem (6.15) which is shown in Figures 6.5 and 6.6. Note

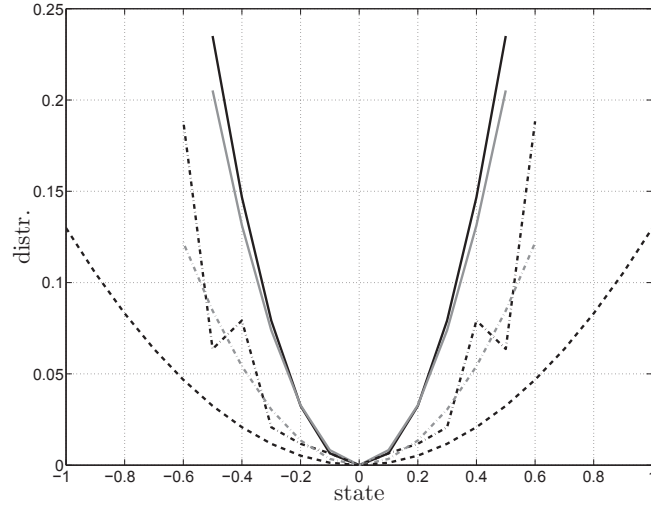


Figure 6.3: Black lines: the initial (dashed line), intermediate (dash-dotted line) and final (solid line) distributions of the agents's states for $\sigma = 0$. Gray lines: the intermediate distribution (dash-dotted line) and final distribution (solid line) resulting from the solution of (6.15) for $\sigma = 0$.

that the distribution computed based on the evolution of the states of the agents matches well with that predicted by the solution to the two-point boundary value problems in all three cases. For completeness, Figure 6.10 shows the time histories of the state of each player when the agents use the control strategies (6.16) for $\sigma = 0$ (top), $\sigma = 0.05$ (middle) and $\sigma = 0.1$ (bottom).

The initial distribution is such that the distribution of agents is relatively low between $x = -0.5$ and $x = -1$. It is therefore expected that the agents move towards this region. This is precisely what occurs, as can be seen in Figure 6.10.

6.5 Conclusion

Mean-field games allows to obtain Nash equilibrium solutions for differential games with infinitely many indistinguishable players. These rely on the solution of two coupled PDEs and can be used to obtain approximate solutions for differential games with a finite number of indistinguishable players. In this chapter it has been demonstrated that for a class of mean-field games, the PDEs characterising the mean-field game boil down to

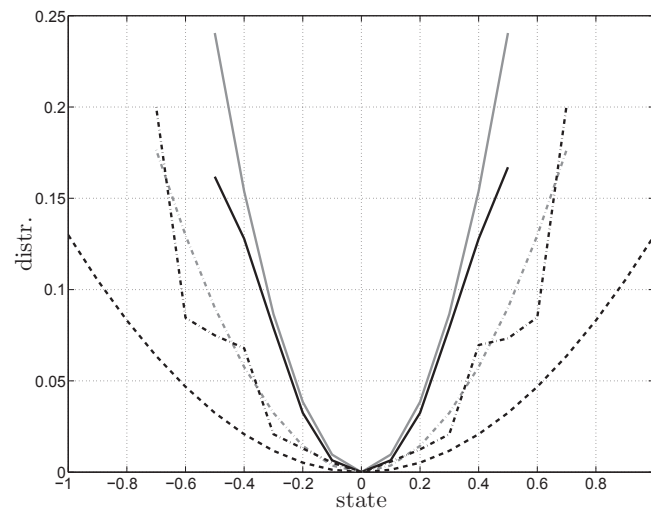


Figure 6.4: Black lines: the initial (dashed line), intermediate (dash-dotted line) and final (solid line) distributions of the agents's states for $\sigma = 0.1$. Gray lines: the intermediate distribution (dash-dotted line) and final distribution (solid line) resulting from the solution of (6.15) for $\sigma = 0.1$.

a system of ODEs with boundary conditions at the initial and final time. Numerical examples illustrate the theory.

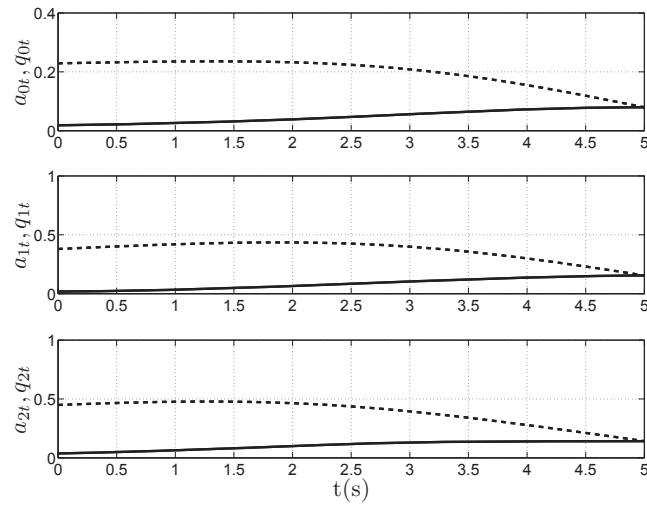


Figure 6.5: Solid lines: time histories of a_{0t} (top) and a_{1t} (middle) and a_{2t} (bottom) for $\sigma = 0.05$. Dashed lines: time histories of q_{0t} (top) and q_{1t} (middle) and q_{2t} (bottom) for $\sigma = 0.05$

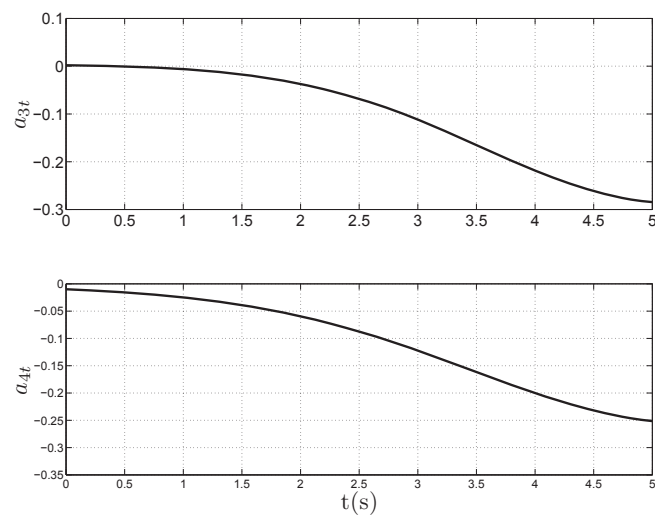


Figure 6.6: Time histories of a_{3t} (top) and a_{4t} (bottom) for $\sigma = 0.05$.

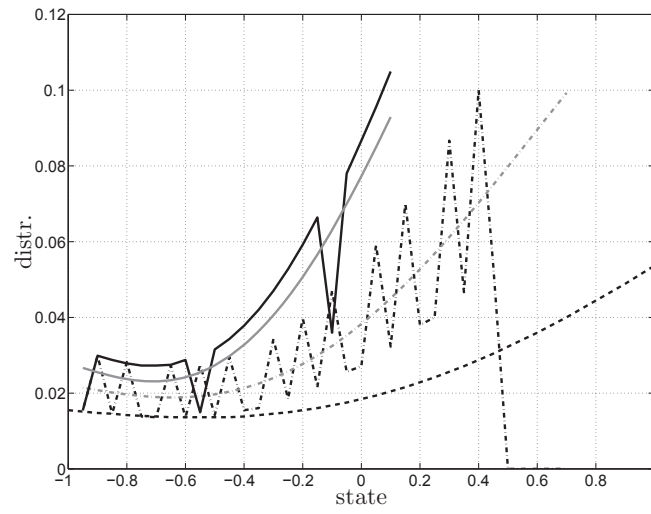


Figure 6.7: Black lines: the initial (dashed line), intermediate (dash-dotted line) and final (solid line) distributions of the agents's states for $\sigma = 0$. Gray lines: the intermediate distribution (dash-dotted line) and final distribution (solid line) resulting from the solution of (6.15) for $\sigma = 0$.

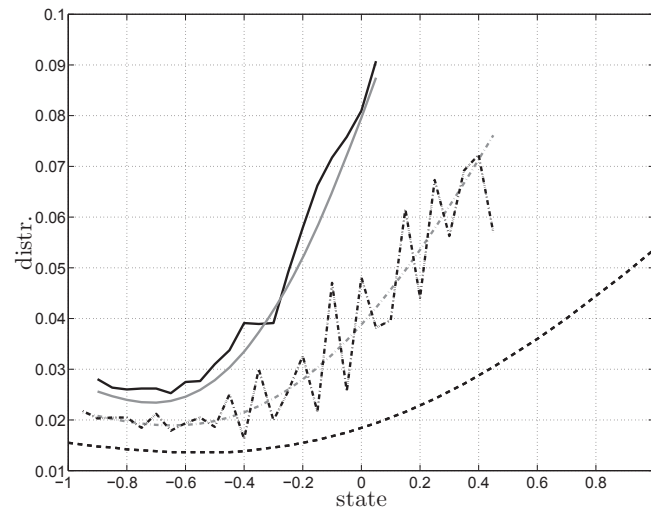


Figure 6.8: Black lines: the initial (dashed line), intermediate (dash-dotted line) and final (solid line) distributions of the agents's states for $\sigma = 0.05$. Gray lines: the intermediate distribution (dash-dotted line) and final distribution (solid line) resulting from the solution of (6.15) for $\sigma = 0.05$.

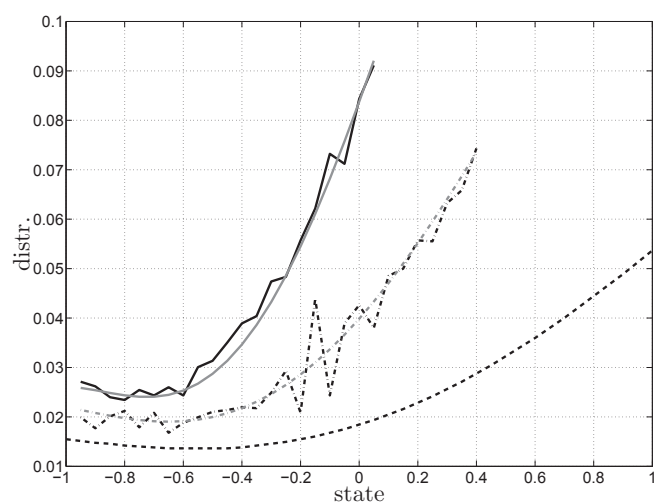


Figure 6.9: Black lines: the initial (dashed line), intermediate (dash-dotted line) and final (solid line) distributions of the agents's states for $\sigma = 0.1$. Gray lines: the intermediate distribution (dash-dotted line) and final distribution (solid line) resulting from the solution of (6.15) for $\sigma = 0.1$.

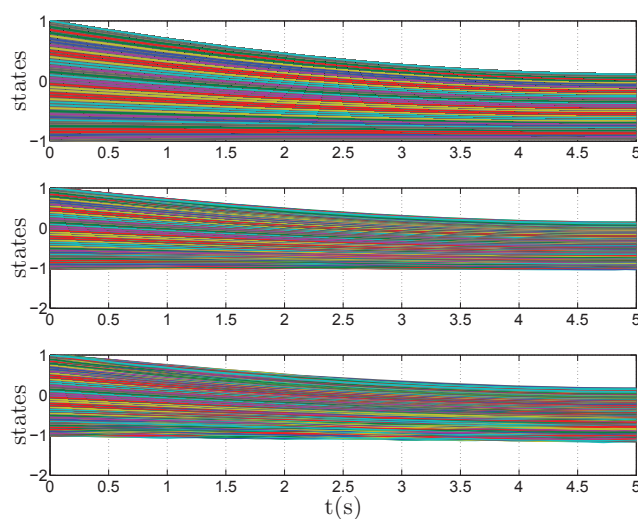


Figure 6.10: Time histories of the state of each player for $\sigma = 0$ (top), $\sigma = 0.05$ (middle) and $\sigma = 0.1$ (bottom).

Chapter 7

Conclusion

7.1 Summary of contributions

The main contributions of the work presented in this thesis are summarised in this chapter before some directions for future research are identified in Section 7.2. The three main parts of the thesis are discussed individually.

Considering the subject of differential game theory, we focus on feedback Nash and Stackelberg equilibrium solutions for a class of nonlinear differential games. The notion of *algebraic $\bar{P}$ matrix solution* is introduced and used to construct dynamic feedback strategies which approximate solutions for the nonlinear differential games. In particular we define the notion of ϵ_α -*equilibrium solutions* and show that the dynamic feedback strategies constitute local ϵ_α -*equilibrium solutions*. The constructive methods rely on a so-called algebraic $\bar{P}$ solution, which is more readily found than solutions to the Hamilton-Jacobi-Isaacs PDEs characterising the differential game and, as a result, the methods allow to study and obtain approximate solutions for problems which would otherwise be difficult to solve. This is illustrated by several numerical examples. Furthermore, as an example of how differential games can be of use when dealing with problems involving multi-agent systems, the *multi-agent collision avoidance problem* is introduced and solved using the constructive methods.

Numerical examples illustrating the theory suggest possible areas of applications

for the developed theory. Among other possible application domains, it is argued that the framework provided by differential game theory can be useful to study problems involving multi-agent systems. This is demonstrated when considering the multi-agent collision avoidance problem and the problem of optimally monitoring a region using a team of mobile agents equipped with sensors. The former problem is solved using the theory developed to obtain approximate solutions for differential games. The latter problem is formulated as a differential game, for which ad-hoc solutions are found assuming the agents satisfy single-integrator dynamics. The resulting trajectories are then interpreted as *trajectory plans* for agents with more complicated unicycle dynamics.

Finally, mean-field games are considered. This is a natural extension of “standard” differential game theory and deals with problems in which a game consists of infinitely many indistinguishable players. It is demonstrated that specific class of mean-field games can be solved by obtaining solutions to a system of ODEs. These can be solved more readily than the original PDEs characterising the mean-field game and consequently the developed theory is a powerful result.

7.2 Future

The results presented herein have the potential for further developments. This thesis is concluded by describing some of these.

The methods introduced to construct approximate feedback solutions for differential games are applicable to certain classes of problems. These results could be extended to a larger variety of problems: two extensions are of particular interest. First, finite-horizon differential games can be considered similarly to what has been done for optimal control problems in [82]. Second, differential games with *non-standard* information structures can be considered, which partly motivates the discussion on approximate solutions for Stackelberg differential games in Chapter 4. In differential games with several players, situations may occur in which the players do not have access to the same information. For example, in [90] the problem in which a differential game consists of two

groups of players is considered. Within each group there is one leader and a Stackelberg game is played between the leader and the rest of the players in the group (*i.e.* the followers), whereas a Nash game is then played between the two groups of players. Clearly, different information structures are possible. Once the problem formulation of differential games with general information structures is formalised the methods for constructing approximate solutions using appropriately modified algebraic $\bar{P}$ matrix solutions can be developed.

Focusing now on the multi-agent problem considered in Chapter 5, it may be possible to achieve a better approximation of the optimal solution for the problem by applying methods similar in spirit to those developed for differential games in Chapter 3. However, to do this, extensions able to deal with finite-horizon problems are necessary. It may also be of interest to pose the monitoring problem as a Stackelberg-type game in which each virtual player is the leader for its corresponding agent, and each agent and virtual player pair play Nash differential games among each other, somewhat similar to the problem considered in [90]. Furthermore, in the current problem formulation it is assumed that each agent has knowledge of the positions of all other agents (and virtual players). Future work includes taking into account communication constraints in the problem formulation. For example, in some applications it may be the case that an agent can only communicate with neighbouring agents. These communication constraints could be considered using graph theory as in [16].

In Chapter 6 a class of mean-field games in which the dynamics are linear, the value function is quadratic and the density distribution is a sum of polynomials is considered. A more general class of problems could be considered. In particular it is of interest to consider problems with nonlinear dynamics and more general value functions and to develop methods for constructing approximate solutions using ideas similar to those introduced in Chapter 3 for differential games.

Bibliography

- [1] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima, and S. Kawatsuma, "Emergency response to the nuclear accident at the Fukushima Daiichi nuclear power plants using mobile rescue robots," *Journal of Field Robotics*, vol. 30, no. 1, pp. 44–63, 2013.
- [2] R. H. Taylor and D. Stoianovici, "Medical robotics in computer-integrated surgery," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 5, pp. 765–781, Oct 2003.
- [3] T. J. Levy and J. D. Beaty, "Revolutionizing prosthetics: Neuroscience framework," *John Hopkins Applied Technical Digest*, vol. 30, no. 3, pp. 223–229, 2011.
- [4] M. M. Bridges, M. P. Para, and M. J. Mashner, "Control system architecture for the modular prosthetic limb," *John Hopkins Applied Technical Digest*, vol. 30, no. 3, pp. 217–222, 2011.
- [5] H. Chang and A. Astolfi, "Control of HIV infection dynamics," *IEEE Control Systems Magazine*, vol. 28, no. 2, pp. 28–39, 2008.
- [6] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. Academic Press.
- [7] M. D. Davis, *Game Theory: A Nontechnical Introduction*. Basic Books, 1983.
- [8] M. Osborne, *An Introduction to Game Theory*. Oxford University Press, Incorporated, 2009.

- [9] A. A. Cournot, *Researches Into the Mathematical Principles of the Theory of Wealth*. Macmillan, 1897.
- [10] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286 – 295, September 1951.
- [11] M. Simaan and J. B. Cruz, Jr., "On the Stackelberg strategy in nonzero-sum games," *Journal of Optimization Theory and Applications*, vol. 11, no. 5, pp. 533–555, 1973.
- [12] J. Milnor, "A nobel prize for John Nash," *The Mathematical Intelligencer*, vol. 17, no. 3, pp. 11–17, 1995.
- [13] T. Basar and P. Bernhard, *H-Infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd ed. Springer, 2008.
- [14] D. J. N. Limebeer, B. D. O. Anderson, and B. Hendel, "A Nash game approach to mixed H_2/H_∞ control," *IEEE Transactions on Automatic Control*, vol. 39, no. 1, pp. 69–82, 1994.
- [15] A. Astolfi and P. Colaneri, "Trading robustness with optimality in nonlinear control," in *American Control Conference*, vol. 6, 2000, pp. 3878–3882 vol.6.
- [16] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, 1st ed. Princeton University Press, 2010.
- [17] B. Nabet, N. E. Leonard, I. Couzin, and S. Levin, "Dynamics and decision making in animal group motion," *Journal of Nonlinear Science*, vol. 19, pp. 399–495, 2009.
- [18] D. A. Paley, N. E. Leonard, R. Sepulchre, D. Grünbaum, and J. K. Parrish, "Oscillator models and collective motion," *IEEE Control Systems Magazine*, pp. 89–105, 2007.
- [19] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das, *Cooperative Control of Multi-Agent Systems: Optimal and Adaptive Design Approaches*. Springer, 2014.
- [20] R. Bachmayer and N. E. Leonard, "Vehicle networks for gradient descent in a sampled environment," in *41st IEEE Conference on Decision and Control*.

- [21] F. Lekien and N. E. Leonard, "Nonuniform coverage and cartograms," *SIAM Journal of Control and Optimisation*, vol. 48, pp. 351–372, 2009.
- [22] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 947–951, 2001.
- [23] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 293–307, 2009.
- [24] T. Nanayakkara, S. F., and M. Jamshidi, *Intelligent Control Systems with an Introduction to System of Systems Engineering*. CRC Press Taylor & Francis Group, 2010.
- [25] H. Azarnoush, B. Horan, P. Sridhar, A. Madni, and M. Jamshidi, "Towards optimization of a real-world robotic-sensor system of systems," in *World Automation Congress*, 2006, pp. 1–8.
- [26] T. Shaneyfelt, S. Erdogan, A. Maredia, G. Vemuri, B. Pachala, D. Yue, S. Karovalia, M.-Z. Lu, S.-Z. Yang, and C. Gao, "Towards net-centric system of systems robotics in air, sea and land," in *IEEE International Conference on System of Systems Engineering*, 2008, pp. 1–6.
- [27] P. Benavidez, K. Nagothu, A. Ray, T. Shaneyfelt, S. Kota, L. Behera, and M. Jamshidi, "Multi-domain robotic swarm communication system," in *IEEE International Conference on System of Systems Engineering*, June 2008, pp. 1–6.
- [28] J. Lasry and P. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
- [29] M. Huang, C. P.E., and R. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems : Centralized and Nash equilibrium," in *42nd IEEE Conference on Decision and Control*, 2003, pp. 98–103.
- [30] —, "Large population cost-coupled lqg problems with non-uniform agents: individual-mass behaviour and decentralized ϵ -Nash equilibria," *IEEE Transactions on Automatic Control*.

- [31] R. Isaacs, *Differential Games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Dover, 1999.
- [32] A. W. Starr and Y. C. Ho, "Nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 3, pp. 184–206, 1969.
- [33] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. The MIT Press, 1994.
- [34] K. J. Åström, *Introduction to Stochastic Control Theory*. Academic Press, 1970.
- [35] T. Mylvaganam, M. Sassano, and A. Astolfi, "Approximate solutions to a class of nonlinear differential games," in *51th IEEE Conference on Decision and Control*, 2012, pp. 6489 – 6494.
- [36] —, "Approximate solutions to a class of nonlinear differential games using a shared dynamic extension," in *12th European Control Conference*, 2013, pp. 710–715.
- [37] —, "Constructive ϵ -Nash equilibria for nonzero-sum differential games," *Submitted to IEEE Transactions on Automatic Control*, 2013.
- [38] —, "A constructive differential game approach to collision avoidance in multi-agent systems," in *American Control Conference*, 2014.
- [39] T. Mylvaganam and A. Astolfi, "Approximate solutions to a class of nonlinear Stackelberg differential games," in *Submitted to the 53rd IEEE Conference on Decision and Control*, 2014.
- [40] —, "Approximate optimal monitoring," in *14th European Control Conference*, 2014.
- [41] —, "Approximate optimal monitoring: preliminary results," in *American Control Conference, Montreal, Canada*, 2012, pp. 4745–4750.
- [42] D. Bauso, T. Mylvaganam, and A. Astolfi, "A two-point boundary value formulation of a mean-field crowd-averse game," in *The 19th World Congress of the International Federation of Automatic Control*, 2014.

- [43] T. Mylvaganam, D. Bauso, and A. Astolfi, "Mean-field games and two-point boundary value problems," in *Submitted to the 53rd Conference of Decision and Control*, 2014.
- [44] D. Bauso, T. Mylvaganam, and A. Astolfi, "Approximate solutions for crowd-averse robust mean-field games," in *The 13th European Control Conference*, 2014.
- [45] E. Dockner, N. Jørgensen, S. Van Long, and G. Sorger, *Differential games in economics and management science*. Cambridge University Press, 2000.
- [46] E. Dockner, G. Feichtinger, and A. Mehlmann, "Noncooperative solutions for a differential game model of fishery," *Journal of Economic Dynamics and Control*, vol. 13, no. 1, pp. 1 – 20, 1989.
- [47] R. D. Fischer and L. J. Mirman, "Strategic dynamic interaction: Fish wars," *Journal of Economic Dynamics and Control*, vol. 16, no. 2, pp. 267 – 287, 1992.
- [48] I. N. Katz, H. Mukai, H. Schättler, M. Zhang, and M. Xu, "Solution of a differential game formulation of military air operations by the method of characteristics," *Journal of Optimization Theory and Applications*, vol. 125, no. 1, pp. 113–135, 2005.
- [49] G. Neumann and S. Schuster, "Continuous model for the rock–scissors–paper game between bacteriocin producing bacteria," *Journal of Mathematical Biology*, vol. 54, no. 6, pp. 815–846, 2007.
- [50] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, pp. 479–519, 2003.
- [51] M. E. Hibbing, C. Fuqua, M. R. Parsek, and S. Brook Peterson, "Bacterial competition: surviving and thriving in the microbial jungle," *Nature Reviews Microbiology*, vol. 8, no. 1, pp. 15–25, 2010.
- [52] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice Hall, 1996.
- [53] K. Zhou, J. C. Doyle, and K. Glover, *Robust Optimal Control*, 1st ed. Prentice Hall, 1997.

- [54] G. P. Papavassilopoulos and M. G. Safonov, "Robust control design via game theoretic methods," in *Proceedings of the 28th IEEE Conference on Decision and Control*, vol. 1, 1989, pp. 382–387.
- [55] D. J. N. Limebeer, B. D. O. Anderson, and B. Hendel, "Nash games and mixed H_2/H_∞ control," in *Robust Control*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 1992, vol. 183, pp. 210–217.
- [56] R. Vinter, *Optimal Control*. Birkhäuser, 2000.
- [57] H. Geering, *Optimal Control with Engineering Applications*. Springer, 2007.
- [58] Y. C. Ho, "Differential games, dynamic optimization, and generalized control theory," *Journal of Optimization Theory and Applications*, vol. 6, pp. 179–209, 1970.
- [59] L. Pontryagin, "On the theory of differential games," *Journal of Russian Mathematical Surveys*, vol. 21, pp. 219–274, 1966.
- [60] Y. C. Ho, A. Bryson, and S. Baron, "Differential games and optimal pursuit-evasion strategies," *IEEE Transactions on Automatic Control*, vol. 10, no. 4, pp. 385 – 389, 1965.
- [61] A. W. Starr and Y. C. Ho, "Further properties of nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 4, pp. 207–219, 1969.
- [62] D. Fudenberg and A. Banerjee, "Word of mouth learning," *Games and Economic Behavior*, vol. 46, pp. 1–22, 2004.
- [63] M. Sassano and A. Astolfi, "Dynamic approximate solutions of the HJ inequality and of the HJB equation for input-affine nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 57, pp. 2490–2503, 2012.
- [64] J. C. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.
- [65] A. J. T. M. Weeren, J. M. Schumacher, and J. C. 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.
- [66] G. P. Papavassilopoulos, J. V. Medanic, and J. B. Cruz, Jr., "On the existence of Nash strategies and solutions to coupled Riccati equations in linear-quadratic games," *Journal of Optimization Theory and Applications*, vol. 28, pp. 49–76, 1979.
- [67] D. Lukes and D. Russell, "A global theory for linear quadratic differential games," *Journal of Mathematical Analysis and Applications*, vol. 33, no. 1, pp. 96–123, 1971.
- [68] G. Freiling, G. Jank, and H. Abou-Kandil, "On global existence of solutions to coupled matrix Riccati equations in closed-loop Nash games," *IEEE Transactions on Automatic Control*, vol. 41, no. 2, pp. 264–269, 1996.
- [69] G. P. Papavassilopoulos and G. J. Olsder, "On the linear-quadratic, closed-loop no-memory Nash game," *Journal of Optimization Theory and Applications*, vol. 42, no. 4, pp. 551–560, 1979.
- [70] J. B. Cruz and C. I. Chen, "Series Nash solution of two-person, nonzero-sum, linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 7, pp. 240–257, 1971.
- [71] D. L. Lukes, "Equilibrium feedback control in linear games with quadratic costs," *SIAM Journal on Control*, vol. 9, no. 2, pp. 234–252, 1971.
- [72] J. C. Engwerda, "Feedback Nash equilibria in the scalar infinite horizon LQ-game," *Automatica*, vol. 36, no. 1, pp. 135–139, 2000.
- [73] —, "The solution set of the n -player scalar feedback nash algebraic riccati equations," *IEEE Transactions on Automatic Control*, vol. 45, no. 12, pp. 2363–2368, 2000.
- [74] N. J. Krikelis and Z. V. Rekasius, "On the solution of the optimal linear control problems under conflict of interest," *IEEE Transactions on Automatic Control*, vol. 16, no. 2, pp. 140–147, 1971.
- [75] J. C. Engwerda, "Feedback Nash equilibria in the scalar infinite horizon LQ-game," *Automatica*, vol. 36, no. 1, pp. 135–139, 2000.

- [76] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou, "The complexity of computing a Nash equilibrium," *SIAM Journal on Computing*, vol. 39, no. 1, pp. 195–259, 2009.
- [77] A. Wernli and G. Cook, "Suboptimal control for the nonlinear quadratic regulator problem," *Automatica*, vol. 11, no. 1, pp. 75–84, 1975.
- [78] J. R. Cloutier, D. Stansbery, and M. Sznaier, "On the recoverability of nonlinear state feedback laws by extended linearization control techniques," in *American Control Conference*, vol. 3, 1999, pp. 1515–1519.
- [79] J. C. Engwerda, "Uniqueness conditions for the infinite-planning horizon open-loop linear quadratic differential game." in *44th IEEE Conference on Decision and Control and European Control Conference*, 2005, pp. 3507 – 3512.
- [80] 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.
- [81] C. Scherer, "The riccati inequality and state-space h_∞ -optimal control," *Univ. Wnrzburg, Germany*, 1990.
- [82] M. Sassano and A. Astolfi, "Approximate finite-horizon optimal control without PDEs," *Systems and Control Letters*, vol. 62, no. 2, pp. 97–103, 2013.
- [83] T. Çimen, "State-dependent Riccati equation (SDRE) control: A survey," in *Proceedings of the 17th World Congress of the International Federation of Automatic Control*, 2008, pp. 3761–3775.
- [84] S. Elloumi, I. Sansa, and N. Braiek, "On the stability of optimal controlled systems with SDRE approach," in *9th International Multi-Conference on Systems, Signals and Devices*, 2012, pp. 1–5.
- [85] K. M. Anstreicer and M. H. Wright, "A note on the augmented Hessian when the reduced Hessian is semidefinite," *SIAM Journal on Optimization*, vol. 11, no. 1, pp. 243–253, 2000.

- [86] M. L. Zeeman, "Extinction in competitive Lotka-Volterra systems," *Proceedings of the American Mathematical Society*, vol. 123, no. 1, pp. 87–96, 1995.
- [87] I. I. Hussein, D. M. Stipanovic, and Y. Wang, "Reliable coverage control using heterogeneous vehicles," in *46th IEEE Conference on Decision and Control*, 2007, pp. 6142–6147.
- [88] I. I. Hussein and D. M. Stipanovic, "Effective coverage control for mobile sensor networks with guaranteed collision avoidance," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 4, pp. 642–657, 2007.
- [89] A. Bensoussan, S. Chen, and S. P. Sethi, "The maximum principle for global solutions of stochastic Stackelberg differential games," *submitted to SIAM Journal on Control and Optimization*, 2012.
- [90] M. Simaan and J. B. Cruz, Jr., "Additional aspects of the Stackelberg strategy in nonzero-sum games," *Journal of Optimization Theory and Applications*, vol. 11, no. 6, pp. 613–626, 1973.
- [91] A. Bensoussan, S. Chen, and S. Sethi, "Feedback Stackelberg solutions of infinite-horizon stochastic differential games," in *Models and Methods in Economics and Management Science*. Springer International Publishing, 2014, vol. 198, pp. 3–15.
- [92] S. J. Rubio, "On the coincidence of the feedback Nash and Stackelberg equilibria in economic applications of differential games," *Economic Working Papers at Centro de Estudios Andaluces*, 2003.
- [93] J. Maciejowski, *Predictive Control: with Constraints*. Pearson Education, 2002.
- [94] J. Rawlings and D. Mayne, *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2009.
- [95] Y. H. Chang, C. Tomlin, and K. Hedrick, "Biologically-inspired coordination of multiple UAVs using sliding mode control," in *American Control Conference*, 2011, pp. 4123–4128.

- [96] P. Panyakeow and M. Mesbahi, "Decentralized deconfliction algorithms for unicycle UAVs," in *American Control Conference*, 2010, pp. 794–799.
- [97] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–355, 2004.
- [98] A. Okabe and A. Suzuki, "Locational optimization problems solved through Voronoi diagrams," *European Journal of Operational Research*, vol. 98, no. 3, pp. 445–456, 1997.
- [99] P. Sujit and D. Ghose, "Multiple agent search of an unknown environment using game theoretical models," in *American Control Conference*, 2004, pp. 5564 – 5569.
- [100] V. Gazi and K. M. Passino, *Swarm Stability and Optimization*, 1st ed. Springer, 2011.
- [101] C. Schroeter and H.-M. Gross, "A sensor-independent approach to RBPF SLAM - Map Match SLAM applied to visual mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 2078 –2083.
- [102] H. Lategahn, A. Geiger, and B. Kitt, "Visual SLAM for autonomous ground vehicles," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1732 –1737.
- [103] D. Castelvechi, "Invasion of the drones: Unmanned aircraft take off in polar exploration," *Scientific American*, pp. 293–307, March 2010.
- [104] J. Wu, S. Yuan, S. Ji, G. Zhou, Y. Wang, and Z. Wang, "Multi-agent system design and evaluation for collaborative wireless sensor network in large structure health monitoring," *Expert Systems with Applications*, vol. 37, no. 3, pp. 2028 – 2036, 2010.
- [105] L. P. Koh and S. A. Wich, "Dawn of drone ecology: low-cost autonomous aerial vehicles for conservation," *Tropical Conservation Science*, vol. 5, no. 2, pp. 121 – 132, 2012.

- [106] R. Vidal, O. Shakernia, H. Kim, D. Shim, and S. Sastry, "Probabilistic pursuit-evasion games: theory, implementation, and experimental evaluation," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 662 – 669, 2002.
- [107] P. Sirdhar, A. Madni, and M. Jamshidi, "Multi-criteria decision making and behavior assignment in sensor networks," in *1st Annual IEEE Systems Conference*, April 2007, pp. 1–7.
- [108] in *Paris-Princeton Lectures on Mathematical Finance*, 2011.
- [109] A. Bensoussan, J. Frehse, and P. Yam, *Mean Field Games and Mean Field Type Control Theory*. Springer, 2013.
- [110] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Networks and Heterogeneous Media*, vol. 7, no. 2, pp. 243–261, 2012.
- [111] H. Tembine, D. Bauso, and T. Basar, "Robust linear quadratic mean-field games in crowd-seeking social networks," in *52nd IEEE Conference on Decision and Control*, 2013, pp. 3134–3139.
- [112] A. Lachapelle and M. T. Wolfram, "On a mean field game approach modeling congestion and aversion in pedestrian crowds," *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1572 – 1589, 2011.
- [113] Loparo, V. Zaitsev, and A. Moussiaux, *The Control Handbook*. CRC Press, 1996.