Title
Nash equilibrium seeking in noncooperative games and multi-agent deployment to planar curves

Permalink
https://escholarship.org/uc/item/9m1408s8

Author
Frihauf, Paul A.

Publication Date
2012

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Nash Equilibrium Seeking in Noncooperative Games and Multi-Agent Deployment to Planar Curves

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

by

Paul A. Frihauf

Committee in charge:
Professor Miroslav Krstic, Chair
Professor Jorge Cortés
Professor Ivana Komunjer
Professor Sonia Martínez
Professor Ruth Williams

2012
The dissertation of Paul A. Frihauf is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____________________________________________________

_____________________________________________________

_____________________________________________________

_____________________________________________________

_____________________________________________________

Chair

University of California, San Diego

2012
DEDICATION

To Jen.
# TABLE OF CONTENTS

Signature Page .......................................................... iii

Dedication ............................................................... iv

Table of Contents ....................................................... v

List of Figures ........................................................ viii

List of Tables .......................................................... xii

Acknowledgements ...................................................... xiii

Vita and Publications ................................................ xvi

Abstract of the Dissertation ........................................ xviii

Chapter 1  Introduction ................................................ 1
  1.1 Nash Equilibrium Seeking in Noncooperative Games ....... 2
  1.2 Multi-agent Deployment to Planar Curves: A PDE-based Approach .... 4
  1.3 Thesis Overview ............................................... 7

Chapter 2  Nash Equilibrium Seeking with Quadratic Payoffs .... 9
  2.1 Two-Player Game .............................................. 10
  2.2 N-Player Games with Quadratic Payoff Functions ......... 14
    2.2.1 General Quadratic Games ............................ 14
    2.2.2 Symmetric Quadratic Games ....................... 18
  2.3 Examples ..................................................... 19
    2.3.1 Duopoly Price Game ................................. 19
    2.3.2 Oligopoly Price Game ............................... 20
  2.4 Approximate Best Response Strategy ....................... 22
    2.4.1 N-Player Games with Stubborn Players ............ 23
    2.4.2 Newton-Based Nash Seeking Strategy .............. 26
    2.4.3 N-Player Games with Gradient-Based and Newton-Based Nash Seeking Players .... 31
  2.5 Summary ..................................................... 33

Chapter 3  Nash Equilibrium Seeking with Non-Quadratic Payoffs .... 34
  3.1 N-Player Games with Non-Quadratic Payoff Functions .... 35
    3.1.1 Averaging Analysis .................................. 37
    3.1.2 Singular Perturbation Analysis .................... 40
  3.2 Example ....................................................... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1</td>
<td>Nash Seeking in the Interior</td>
<td>44</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Nash Seeking with Projection</td>
<td>48</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary</td>
<td>50</td>
</tr>
<tr>
<td>4.1</td>
<td>Nash Equilibrium Seeking</td>
<td>52</td>
</tr>
<tr>
<td>4.2</td>
<td>Convergence Results</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Uncountably-Many Player Price Game</td>
<td>58</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary</td>
<td>61</td>
</tr>
<tr>
<td>5.1</td>
<td>Problem Statement</td>
<td>63</td>
</tr>
<tr>
<td>5.2</td>
<td>Optimal Control via Extremum Seeking</td>
<td>66</td>
</tr>
<tr>
<td>5.3</td>
<td>Error System</td>
<td>67</td>
</tr>
<tr>
<td>5.4</td>
<td>Convergence Result</td>
<td>71</td>
</tr>
<tr>
<td>5.5</td>
<td>Simulation Results</td>
<td>76</td>
</tr>
<tr>
<td>5.5.1</td>
<td>Example 1: Scalar System</td>
<td>76</td>
</tr>
<tr>
<td>5.5.2</td>
<td>Example 2: Unstable Third-Order System</td>
<td>77</td>
</tr>
<tr>
<td>5.6</td>
<td>Summary</td>
<td>80</td>
</tr>
<tr>
<td>6.1</td>
<td>Leader-Enabled Deployment</td>
<td>82</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Decoupled 1-D Deployments</td>
<td>84</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Complex-valued 2-D Deployment</td>
<td>88</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Design Procedure for Desired Deployment Profiles</td>
<td>91</td>
</tr>
<tr>
<td>6.2</td>
<td>Leader Feedback Design</td>
<td>91</td>
</tr>
<tr>
<td>6.3</td>
<td>Closed-Loop Stability</td>
<td>95</td>
</tr>
<tr>
<td>6.4</td>
<td>Leader Observer Design</td>
<td>97</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Leader, Neighbor, and Anchor Measurements</td>
<td>98</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Leader and Neighbor Measurements</td>
<td>100</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Output Feedback</td>
<td>101</td>
</tr>
<tr>
<td>6.5</td>
<td>Discretized Agent Control Laws</td>
<td>101</td>
</tr>
<tr>
<td>6.6</td>
<td>Simulations</td>
<td>105</td>
</tr>
<tr>
<td>6.7</td>
<td>Summary</td>
<td>111</td>
</tr>
<tr>
<td>7.1</td>
<td>Deployment to Circular Arcs</td>
<td>114</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Nonlinear PDE Model</td>
<td>115</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Convergence Result</td>
<td>116</td>
</tr>
<tr>
<td>7.2</td>
<td>Deployment to Elliptical Arcs</td>
<td>116</td>
</tr>
<tr>
<td>7.3</td>
<td>Discretized Agent Control Laws</td>
<td>123</td>
</tr>
<tr>
<td>7.4</td>
<td>Simulations</td>
<td>125</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1: Topical relationship among the chapters. Dashed lines indicate an indirect relationship. ................................. 8

Figure 2.1: Deterministic Nash seeking schemes applied by players in a duopoly market structure. ..................................... 12

Figure 2.2: The (a) price, (b) sales, and (c) profit time histories for P1 and P2 when implementing the Nash seeking scheme (2.9)–(2.10). The dashed lines indicate the values at the Nash equilibrium. ... 13

Figure 2.3: A model of sales \(s_1, s_2, s_3\) in a three-firm oligopoly with prices \(u_1, u_2, u_3\) and total consumer demand \(S_d\). The desirability of product \(i\) is proportional to \(1/R_i\)................................. 21

Figure 2.4: The (a) price, (b) sales, and (c) profit time histories of firms P1, P2, P3, and P4 when implementing the Nash seeking strategy (2.9)–(2.10). The dashed lines indicate the values at the Nash equilibrium. .............................................. 23

Figure 2.5: The (a) price, (b) sales, and (c) profit time histories of firms P1, P2, P3, and P4 when P1 is a stubborn player and P2, P3, and P4 implement the Nash seeking strategy (2.9)–(2.10). The dashed lines indicate the steady-state best response values of players P2, P3, and P4 given the constant price of P1. .... 27

Figure 3.1: The steady-state payoff function surfaces (3.27) with both their associated reaction curves (black) (3.28)–(3.29), which lie in the action set \(U\), and the extremals \(\partial J_i/\partial u_i = 0\) (dashed yellow), which lie outside of \(U\), superimposed. ................................. 44

Figure 3.2: (a) The reaction curves associated with the steady-state payoff functions (3.27) with the stable Nash equilibrium (green circle), the unstable Nash equilibrium (red square), and the Nash equilibrium on the boundary (blue star), and (b) the convergence bias (+) relative to each interior Nash equilibrium, due to the non-quadratic payoff function of player 2 and lies along the reaction curve of player 1. ................................. 45

Figure 3.3: Time history of the two-player game initialized at (a) \((u_1, u_2) = (0.2, 0.9)\), and (b) \((u_1, u_2) = (0.1, 0.15)\). ................................. 46
Figure 3.4: Phase portrait of the \( \hat{u} \)-reduced model average system, with the players' \( \hat{u} \)-trajectories superimposed for each initial condition. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and Nash equilibrium on \( \partial U \) (blue star) are denoted. The shaded region is the set difference \( U \setminus \hat{U} \). If \( \hat{u}_i \in \hat{U} \), then \( u_i \in U \). The zoomed-in area in shows the players' convergence to an almost-periodic orbit that is biased away from the stable Nash equilibrium and centered on the reaction curve \( l_1 \). The average actions of the players are marked by \( \times \).

Figure 3.5: (a) Time history of the two-player game initialized at \((0.2, 0.03)\) when implementing the modified Nash seeking strategy \((3.33)-(3.34)\), and (b) planar \( \hat{u} \)-trajectories superimposed on the \( \hat{u} \)-reduced model average system phase portrait. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and the Nash equilibrium on \( \partial U \) (blue star) are denoted. The shaded region is the set difference \( U \setminus \hat{U} \).

Figure 3.6: (a) Time history of the two-player game initialized at \((0.2, 0.03)\) when implementing the modified Nash seeking strategy \((3.38)-(3.39)\), and (b) planar \( \hat{u} \)-trajectories superimposed on the \( \hat{u} \)-reduced model average system phase portrait. The unstable Nash equilibrium (red square), its eigenvectors (magenta arrows), and the Nash equilibrium on \( \partial U \) (blue star) are denoted. The shaded region is the set difference \( U \setminus \hat{U} \).

Figure 4.1: Deterministic extremum seeking scheme employed by uncountably-many players with the Nash seeking loop for player \( x \in [0, 1] \) depicted.

Figure 4.2: Price evolution of the uncountably-many player oligopoly price game.

Figure 4.3: Nash equilibrium \( u^*(x) \) (dashed) of the oligopoly price game with the players' price (blue) at \( t = 400 \) sec superimposed.

Figure 4.4: Payoff evolution of the uncountably-many player oligopoly price game.

Figure 5.1: Extremum seeking scheme employed to find the optimal control sequence \( \{u^*_k\} \) that minimizes the cost function \( J(u) \). The \( k \)-th controller loop is depicted with the control values \( \{u_i\}_{i \neq k} \) representing the output of the extremum seeking loops not shown.

Figure 5.2: Condition number of \( \kappa(H) \) versus \( \alpha \) for various time horizons \( N \) with a log scale for the y-axis (top) and a linear scale (bottom).
Figure 5.3: Eigenvalues of $H$ versus $\alpha$ for $N = 5$ (top) and $N = 10$ (bottom). The maximum eigenvalue is approximately the condition number $\kappa(H) = \lambda_{\max}(H)/\lambda_{\min}(H)$, shown in Figure 5.2. 

Figure 5.4: Trajectory of $u_0$ and $u_1$ for Example 2 for the constrained (blue) and unconstrained (red) cases. The cost level sets are shown with the unconstrained minimum (green circle). The dashed lines denote the input constrained set with the optimum value on the boundary (cyan square). 

Figure 5.5: Time history of $u_2$ for both the constrained (top) and unconstrained (bottom) cases in Example 2. Black lines denote the optimal values. 

Figure 6.1: Communication topology imposed by spatial discretization. The leader needs global information (dashed) unless it uses output feedback. Legend: triangle = anchor, plus = follower, star = leader. 

Figure 6.2: Lissajous curves for various values of $C, D,$ and $E$ with $A = B = 1$, $\alpha \in [0, 1]$. The dots represent agents. Legend: black = 1 agent, red = 2 agents. 

Figure 6.3: Agents in the ellipse deployment family (a) rendezvous at the origin with (b) the corresponding control effort (top) and observer error (bottom). The agents converge to the origin after the observer error is negligible ($\approx 14$ sec). In (b), we plot the control effort and observer error for $t \in [0, 15]$ sec to highlight the transients since all the curves decay to zero after 15 sec. 

Figure 6.4: Agent trajectories for (a) the ellipse deployment family when deploying to a circle and then to a rotated ellipse, and (b) the Lissajous figure 8 deployment family when deploying to a figure 8 and a distorted figure 8. 

Figure 6.5: Agent trajectories (a) using an advection-diffusion equation and a reaction-diffusion equation for the $x$- and $y$-deployments respectively, and (b) for the complex-valued circle deployment family when deploying to a circle centered at $(0, 0)$ and to another centered at $(4, -4)$. 

Figure 6.6: Agent trajectories for (a) the complex-valued spiral deployment family when deploying to a logarithmic spiral, and (b) a helical deployment with $t \in [0, 30]$ sec. 

Figure 7.1: Communication topology imposed by spatial discretization with arrows denoting the information flow. Legend: triangle = anchor, plus = follower. 

Figure 7.2: Agent trajectories for deployment onto a circular arc with nonuniform agent spacing.
Figure 7.3: Agent trajectories for deployments with nonuniform agent spacing when (a) deploying from a circular arc to a circle formation and (b) redeploying from a circle formation to a spiral deployment.
Table 6.1: Basis functions for 1-D deployment curves of the reaction-advection-diffusion equation. 85
ACKNOWLEDGEMENTS

Many people deserve thanks for helping me reach this point in my academic career, and this small section is, by no means, adequate to express my sincere gratitude. But it is a start.

First, I would like to thank my wife Jen for her constant love and support. Experiencing graduate school together has been far more rewarding than if we had taken turns. And to our beautiful daughter Adelyn, thanks for the extra motivation to finish.

This work would not have happened without the guidance and expertise of my advisor Professor Miroslav Krstic, to whom I am thankful. I would also like to thank Professor Tamer Başar for taking the time and effort to help guide the research of a student who has no direct connection to him, is many miles away, and sometimes sends him last minute drafts of papers.

I would like to thank my committee members for your time, effort, and willingness to fit me into your already crowded schedules.

I would like to extend a sincere thank you my lab mates: Antranik Siranosian and Jennie Cochran for showing me the ropes; Nima Ghods, Nikos Bekiaris-Liberis, and Halil Basturk for making the lab an enjoyable place to work; James Krieger, Gideon Prior, James Gray, and Joe Doblack for the enjoyable study groups; Shu-jun Liu and Beibei Ren for being so nice and working so hard on our papers; Alex Scheinker for occasionally gracing us with his presence when he is not in Los Alamos; and Azad Ghaffari, David Ryan, Mariana Moretti, and Shuxia Tang for gracefully accepting the baton from those of us leaving the lab.

Thank you, Dan Miller and Ben Regner, for the healthy doses of perspective and for sometimes picking up the tab. You guys are good company.

Tyler and Jena Hales, thank you for being such good friends. It’s comforting to know that these past five years are simply the start of our friendship.

I am proud to consider myself an honorary member of the Neuroscience (and friends) program. I have been lucky to be part of this tightly knit group. Thank you, Adam and Maria Koerner, Tristan Shuman, Denise Cai, my fellow honorary member David Lipshutz, Sarah Parylak, Priya Velu, Adam Calhoun,
Dan Knudsen, Matt O’Sullivan, Nick Bevins, and Jeremy Biane (even though you occasionally make us run wind sprints).

Thank you to the Becker and Newton families for making San Diego feel like home.

Of course, I would like to thank my parents. Whenever an opportunity arose for me, you never said “no.” And to my brother, thank you for setting the example of how to earn a Ph.D. and raise a loving family.

The research contained in this dissertation was partially funded by a National Defense Science and Engineering Graduate (NDSEG) Fellowship. I would also like to acknowledge the ARCS Foundation (Achievement Rewards for College Scientists) for additional financial support.

This dissertation includes reprints or adaptations of the following papers:

- P. Frihauf, M. Krstic, and T. Başar, “Nash equilibrium seeking in noncooperative games,” *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1192–1207, 2012. (Chapters 2 and 3, excluding Sections 2.4.2 and 2.4.3)


The dissertation author is the primary investigator and author of these works.
VITA

<table>
<thead>
<tr>
<th>Year</th>
<th>Experience Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Industrial Engineering Intern, The Boeing Company, Seattle, WA</td>
</tr>
<tr>
<td>2004</td>
<td>Summer Research Program, MIT Lincoln Laboratory, Lexington, MA</td>
</tr>
<tr>
<td>2004–2005</td>
<td>Research Assistant, Microflight Corporation, St. Louis, MO</td>
</tr>
<tr>
<td>2005</td>
<td>B.S./M.S. in Systems Science and Engineering, Washington University in St. Louis, St. Louis, MO</td>
</tr>
<tr>
<td>2005–2007</td>
<td>Control Engineer, Johns Hopkins University Applied Physics Laboratory, Laurel, MD</td>
</tr>
<tr>
<td>2007–2010</td>
<td>NDSEG Research Fellow, University of California, San Diego, La Jolla, CA</td>
</tr>
<tr>
<td>2010</td>
<td>Teaching Assistant, Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA</td>
</tr>
<tr>
<td>2011</td>
<td>Algorithms Intern, Cymer, Inc., San Diego, CA</td>
</tr>
<tr>
<td>2011–2012</td>
<td>Research Assistant, Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA</td>
</tr>
<tr>
<td>2011–2012</td>
<td>Achievement Rewards for College Scientists (ARCS) Foundation Scholar, University of California, San Diego, La Jolla, CA</td>
</tr>
<tr>
<td>2012</td>
<td>Ph.D. in Engineering Sciences (Mechanical Engineering), University of California, San Diego, La Jolla, CA</td>
</tr>
</tbody>
</table>

PUBLICATIONS

Journal Articles


**Articles in Edited Volumes**


**Conference Proceedings**


We present two research topics in multi-agent systems: Nash equilibrium seeking for players in unknown noncooperative games and the deployment of agents onto families of planar curves. The designed controllers either extract or encode the information necessary for the agents to achieve their objectives. Selfish players converge to a Nash equilibrium despite not knowing the game’s mathematical structure, and follower agents achieve deployment to planar curves without explicit knowledge of the desired formation.

To solve static, noncooperative games, we introduce a non-model based learning strategy, based on the extremum seeking approach, for $N$-player and
infinitely-many player games. In classical learning algorithms, each player has knowledge of its payoff’s functional form and of the other players’ actions, whereas in the proposed algorithm, the players measure only their own payoff values. For games with non-quadratic payoffs, convergence to a Nash equilibrium is not perfect, but is biased in proportion to the algorithm’s perturbation amplitudes and the payoff functions’ higher derivatives. In games with infinitely-many players, no single player may affect the outcome of the game, and yet, we show that a player may converge to the Nash equilibrium by measuring only its own payoff value. Inspired by the infinitely-many player case, we also present an extremum seeking approach for locally stable attainment of the optimal open-loop control sequence for unknown discrete-time linear systems, where not even the dimension of the system is known.

To achieve deployment to a family of planar curves, we model the agents’ collective dynamics with the reaction-advection-diffusion class of PDEs. The PDE models, whose state is the position of the agents, incorporate the agents’ feedback laws, which are designed based on a spatial internal model principle. Namely, the agents’ feedback laws allow the agents to deploy to a family of geometric curves that correspond to the model’s equilibrium profiles. Stable deployment is ensured by leader feedback, designed using techniques for the boundary control of PDEs. A nonlinear PDE model is also presented that does not require leader feedback and enables the agents to deploy to planar arc formations.
Chapter 1

Introduction

This work considers two topics in multi-agent systems: Nash equilibrium seeking in noncooperative games and deployment onto families of planar curves. In both topics, the information contained in an agent’s controller is central to its ability to meet its objective, whether that objective is selfish or cooperative in nature. Using a Nash equilibrium seeking strategy, players in an unknown noncooperative game converge to a Nash equilibrium by estimating, on the average, the gradient of their respective payoff functions. While the structure of the game is unknown, the payoff values provide enough information for each player to achieve convergence. Using controllers developed from PDE models, agents are able to deploy to families of planar curves while using a single controller that does not possess explicit information of the desired formation. The encoding of the feasible planar deployments into the agents’ controllers represents a spatial application of the internal model principle [35]. These formations are stabilized by a leader agent, whose feedback control is designed using PDE boundary control techniques.

The following sections provide a more thorough introduction to these two topics.
1.1 Nash Equilibrium Seeking in Noncooperative Games

The development of algorithms to achieve convergence to a Nash equilibrium has been a focus of researchers for several decades. Advances in both theory and technology have helped continue this line of research as game theory has found applications across a wide-array of disciplines. We study the problem of solving noncooperative games in real time by employing a non-model based approach where the players determine their actions using only their own measured payoff values. By utilizing deterministic extremum seeking with sinusoidal perturbations, the players attain their Nash strategies without the need for any model information. Games with \( N \) players and with infinitely-many players are considered.

For games with \( N \) players, we analyze both static games and games with dynamics, where the players’ actions serve as inputs to a general, stable nonlinear dynamic system whose outputs are the players’ payoff values. In the latter scenario, the dynamic system evolves on a faster time scale compared to the time scale of the players’ strategies, resulting in a static game being played at the steady state of the dynamic system. In these scenarios, the players possess either quadratic or non-quadratic payoff functions, which may result in multiple, isolated Nash equilibria. We quantify the convergence bias relative to the Nash equilibria that results from payoff function terms of higher order than quadratic. For games with infinitely-many players, we restrict our analysis to static games with quadratic payoff functions.

Most algorithms designed to achieve convergence to Nash equilibria require modeling information of the game and assume the players can observe the actions of the other players. Two classical examples are the best response and fictitious play strategies, where each player chooses the action that maximizes its payoff given the actions of the other players. The Cournot adjustment, a myopic best response strategy, was first studied by Cournot [29] and refers to the scenario where a firm in a duopoly adjusts its output to maximize its payoff based on the known output of its competitor. The strategy known as fictitious play (employed
in finite games), where a player devises a best response based on the history of
the other players’ actions was introduced in [18] in the context of mixed-strategy
Nash equilibria in matrix games. In [86], a gradient-based learning strategy where
a player updates its action according to the gradient of its payoff function was
developed. Stability of general player adjustments was studied in [66] under the
assumption that a player’s response mapping is a contraction, which is ensured
by a diagonal dominance condition for games with quadratic payoff functions.
Distributed iterative algorithms for the computation of equilibria in a general class
of non-quadratic convex Nash games were designed in [64], and conditions for the
contraction of general nonlinear operators were obtained to achieve convergence.

In more recent work, a dynamic version of fictitious play and gradient re-
response, which also includes an entropy term, is developed in [93] and is shown
to converge to a mixed-strategy Nash equilibrium in cases that previously devel-
oped algorithms did not converge. In [112], a synchronous distributed learning
algorithm, where players remember their own actions and utility values from the
previous two times steps, is shown to converge in probability to a set of restricted
Nash equilibria. An approach that is similar to our Nash seeking method (found
in [38, 58] and in this paper) is studied in [100] to solve coordination problems in
mobile sensor networks. Additional results on learning in games can be found in
[21, 34, 40, 50, 94]. Some diverse engineering applications of game theory include
the design of communication networks in [1, 10, 68, 89, 113], integrated structures
and controls in [85], and distributed consensus protocols in [12, 90, 69]. A com-
prehensive treatment of static and dynamic noncooperative game theory can be
found in [11].

Inspired by the game scenario with infinitely-many players, we also study
the optimal control problem for a discrete-time linear system that is entirely un-
known—only the system’s output is known. The construction of optimal control
signals for dynamical systems is a much-studied area of control theory, and recent
studies have been dedicated to the optimal control problem for unknown systems,
relying mainly on learning and dynamic programming methods. In [77], an ob-
server and neural network are used to approximate a value function, which is used
in a policy iteration scheme to determine an optimal control strategy for an unknown system whose relative degree is known. Policy iteration and value iteration algorithms are derived in [63] that require only output feedback to achieve the optimal control policy for unknown systems of known dimension. In [31], a neural network is used to learn the unknown system model, which is then used to find offline the optimal control policy using approximate dynamic programming. Techniques similar to those presented in this work are used to achieve desired plasma current profiles in a magnetic fusion reactor in [79].

The results of these works extend the extremum seeking method [3, 23, 67, 71, 76, 101, 103, 107], originally developed for standard optimization problems. Many works have used extremum seeking, which performs non-model based gradient estimation, for a variety of applications, such as steering vehicles toward a source in GPS-denied environments [24, 25, 111], optimizing the control of HCCI engines [56] and nonisothermal continuously stirred tank reactors [44], reducing the impact velocity of an electromechanical valve actuator [82], and controlling flow separation [13] and Tokamak plasmas [20].

1.2 Multi-agent Deployment to Planar Curves: A PDE-based Approach

Much research has been conducted in multi-agent formation control, leading to many approaches for stable deployment onto curves. However, the agents in many of these works implement controllers that depend on the desired deployment, hence the parameters of each agent’s controller must be updated to move the agents from one deployment to another, which may be cumbersome for systems with large numbers of agents.

We propose a framework that enables agents to achieve deployment families while employing a single controller with no knowledge of the desired deployment. These families correspond to the potentially unstable, nonzero equilibria of either two decoupled, linear reaction-advection-diffusion PDEs or one complex-valued, linear Ginzburg-Landau PDE and are stabilized by a leader agent. These PDEs
models serve as an application of the internal model principle [35] in a *spatial* sense, as it allows the agents to deploy to a family of deployments derived from the equilibrium profiles. The agents’ positions, needed for the leader’s feedback, are estimated by observers that require position information from only the leader’s nearest neighbor. Thus, when the leader employs output feedback, all the agents use only local information. For large multi-agent systems, our framework allows a user to deploy *many* agents to multiple configurations while communicating with only *two* agents, the leader agent and another agent termed the *anchor*.

We also introduce a nonlinear approach for deployment to a family of planar arcs by using two anchor agents to manipulate the group geometry. The arc formations, parameterized by the desired radius, correspond to the equilibrium profiles of a nonlinear PDE model, which, as in the linear case, is a spatial application of the internal model principle. The agents achieve exponential convergence to these stationary formations, without the need for leader feedback, using only local information, but the convergence rate cannot be arbitrarily set.

This work draws from multi-agent systems research in formation control, estimation, and PDE-related designs. In formation control, feasible geometric patterns are characterized for agents with global information in [102], and stabilization of any geometric pattern using consensus-based controllers, which require each agent to have knowledge of the desired formation, are studied in [32, 65, 60]. Convergence to generalized regular polygons, a point, or a logarithmic spiral pattern is shown for unicycles under cyclic pursuit in [70, 81], a sequence of maneuvers between formation patterns is achieved with a behavior-based approach in [61], and planar formations, such as circle deployments, are controlled in a Lie group setting in [53]. By selecting appropriate density functions that are known by all the agents, coverage control algorithms designed in [28] can be used to achieve deployments onto a desired planar curve and 2-D distributions within a desired planar curve. In [49], decentralized controllers that maintain connectivity are used to form geometric patterns specified by a smooth function.

Other works use a leader agent to influence the collective behavior of the agents. Artificial potentials and virtual leaders are used to control the group’s
geometry and mission in [62]. In leader-follower systems, nonholonomic followers use nonlinear controllers to stabilize their relative distances and orientation in [30], leader-to-formation stability gains are used to quantify a formation’s stability properties in [104], and bounds on a leader’s velocity and the curvature of its path, which guarantee the existence of a follower’s formation-maintaining controller, are determined in [27]. A leader agent is used to steer a formation in [60] and to optimally transfer agents to desired waypoints at specified times in [16]. A group of leaders herd follower agents to a target location while ensuring they remain inside the convex polytope defined by the leaders in [51]. Leader agents in [65, 19] cause a group of follower agents, who employ consensus-based control laws, to deploy onto a line segment. The follower agents do not have knowledge of the formation, which inspires this work, and the linear deployment family corresponds to our simplest control design based on the heat equation.

Multi-agent estimation research has focused mainly on dynamic consensus filters. Vehicles use an information exchange methodology, whose stability is decoupled from the local control of the vehicles, to reach consensus on a formation’s center in [32]. Laplacian consensus dynamics are extended in [99] to handle time-varying signals, arbitrary time delays, and splitting and merging networks. Dynamic consensus filters are developed in [36] and implemented in [110] to estimate global information for use in an agent’s local controller to achieve the system’s desired global behavior. Dynamic consensus filters are also used in [91] to stabilize parallel or circular collective motion, recovering the results in [92]. We design observers using the backstepping approach for PDEs with boundary sensing [97] to estimate the follower agents’ positions for use in the leader agent’s feedback.

Recent years have seen an increase in multi-agent research utilizing PDEs for both design and analysis [9, 14, 52, 54, 83, 88]. In particular, [14] uses a PDE from image processing to design boundary tracking controllers, [83] models a swarm as an incompressible fluid for pattern generation, and [9] studies the stability of large vehicular platoons using a linear hyperbolic PDE. The Partial difference Equation (PdE) framework is used in [33] to show that Laplacian control, analyzed in [74, 78], coincides with the heat equation. It is also used to develop control laws
in [51, 57], where in [57], agents use model reference adaptive control laws to track desired trajectories, using either the heat equation or the wave equation as reference models. By employing boundary control techniques for PDEs, we are able to utilize PDE models that are open-loop unstable, which increases greatly the number of feasible deployment families for the agents.

1.3 Thesis Overview

The remainder of this thesis is outlined below:

Chapter 2 introduces the Nash seeking learning strategy with a duopoly price game example, proves convergence to the Nash equilibrium in $N$-player games with quadratic payoff functions, and shows that this strategy approximates a player’s continuous-time best response dynamics.

Chapter 3 extends the results of Chapter 2 to $N$-player games with non-quadratic payoff functions that are the output of a dynamic system. Convergence to a Nash equilibrium is not perfect, but is biased in proportion to the perturbation amplitudes and the higher derivatives of the payoff functions. We quantify the size of these residual biases.

Chapter 4 extends the Nash seeking strategy to infinitely-many player games with quadratic payoff functions. Even though a single player cannot influence the outcome of the game, we prove that it can still achieve convergence to the Nash equilibrium by measuring only its own payoff value.

Chapter 5 presents finite-horizon LQ optimal control for unknown discrete-time systems using extremum seeking. Convergence to the optimal open-loop control sequence of a repetitive process is obtained using discrete-time extremum seeking.

Chapter 6 introduces leader-enabled multi-agent deployment to families of planar curves. PDE models encode the deployment families into the agents’ controllers, allowing the follower agents to stabilize multiple deployments without changing their control laws. Stability is ensured by leader feedback, which is designed using PDE boundary control techniques.
Figure 1.1: Topical relationship among the chapters. Dashed lines indicate an indirect relationship.

Chapter 7 extends the results of Chapter 6 to a nonlinear PDE model for deployment to planar arcs. This family of deployments does not require leader feedback for stabilization, but the convergence rate of the agents can no longer be arbitrarily set.

Figure 1.1 provides a topical relationship among the chapters. Chapter 4 can be viewed as a single-player dynamic game with a quadratic payoff function. Chapters 5, 6, and 7 model the infinitely-many agents as a continuum.
Chapter 2

Nash Equilibrium Seeking with Quadratic Payoffs

We develop a Nash seeking strategy to stably attain a Nash equilibrium in noncooperative \( N \)-player games. The key feature of our approach is that the players do not need to know the mathematical model of their payoff functions or of the underlying model of the game. They only need to measure their own payoff values when determining their respective time-varying actions, which classifies this learning strategy as *radically uncoupled* according to the terminology of [34]. When the players have quadratic payoff functions, this strategy is semi-globally practically asymptotically stable, but we focus on local convergence results in anticipation of the extension to games with non-quadratic payoff functions in Chapter 3.

This chapter is organized as follows: We motivate our Nash seeking strategy with a duopoly price game in Section 2.1, prove convergence to the Nash equilibrium in \( N \)-player games with quadratic payoff functions in Section 2.2, and provide numerical examples in Section 2.3. Section 2.4 shows that this strategy approximates a player’s best response dynamics and introduces the Newton-based Nash equilibrium seeking approach. Section 2.5 summarizes the results.
2.1 Two-Player Game

To introduce our Nash seeking algorithm, we first consider a specific two-player noncooperative game, which for example, may represent two firms competing for profit in a duopoly market structure. Common duopoly examples include the soft drink companies, Coca-Cola and Pepsi, and the commercial aircraft companies, Boeing and Airbus. We present a duopoly price game in this section for motivational purposes before proving convergence to the Nash equilibrium when $N$ players with quadratic payoff functions employ our Nash seeking strategy in Section 2.2.

Let players P1 and P2 represent two firms that produce the same good, have dominant control over a market, and compete for profit by setting their prices $u_1$ and $u_2$, respectively. The profit of each firm is the product of the number of units sold and the profit per unit, which is the difference between the sale price and the marginal or manufacturing cost of the product. In mathematical terms, the profits are modeled by

$$J_i(t) = s_i(t) (u_i(t) - m_i),$$  \hspace{1cm} (2.1)$$

where $s_i$ is the number of sales, $m_i$ the marginal cost, and $i \in \{1, 2\}$ for P1 and P2. Intuitively, the profit of each firm will be low if it either sets the price very low, since the profit per unit sold will be low, or if it sets the price too high, since then consumers will buy the other firm’s product. The maximum profit is to be expected to lie somewhere in the middle of the price range, and it crucially depends on the price level set by the other firm.

To model the market behavior, we assume a simple, but quite realistic model, where for whatever reason, the consumer prefers the product of P1, but is willing to buy the product of P2 if its price $u_2$ is sufficiently lower than the price $u_1$. Hence, we model the sales for each firm as

$$s_1(t) = S_d - s_2(t),$$  \hspace{1cm} (2.2)$$

$$s_2(t) = \frac{1}{p} (u_1(t) - u_2(t)).$$  \hspace{1cm} (2.3)$$
where the total consumer demand $S_d$ is held fixed for simplicity, the preference of the consumer for P1 is quantified by $p > 0$, and the inequalities $u_1 > u_2$ and $(u_1 - u_2)/p < S_d$ are assumed to hold.

Substituting (2.2), (2.3) into (2.1) yields expressions for the profits $J_1(u_1, u_2)$ and $J_2(u_1, u_2)$ that are both quadratic functions of the prices $u_1$ and $u_2$, namely,

$$J_1 = \frac{-u_1^2 + u_1 u_2 + (m_1 + Sdp)u_1 - m_1 u_2 - Sdp m_1}{p},$$  \hfill (2.4) \\
$$J_2 = \frac{-u_2^2 + u_1 u_2 - m_2 u_1 + m_2 u_2}{p},$$  \hfill (2.5)

and thus, the Nash equilibrium is easily determined to be

$$u_1^* = \frac{1}{3}(2m_1 + m_2 + 2Sdp),$$  \hfill (2.6) \\
$$u_2^* = \frac{1}{3}(m_1 + 2m_2 + Sdp).$$  \hfill (2.7)

To make sure the constraints $u_1 > u_2$, $(u_1 - u_2)/p < S_d$ are satisfied by the Nash equilibrium, we assume that $m_1 - m_2$ lies in the interval $(-Sdp, 2Sdp)$. If $m_1 = m_2$, this condition is automatically satisfied.

For completeness, we provide here the definition of a Nash equilibrium $u^* = [u_1^*, \ldots, u_N^*]^T$ in an $N$-player game:

$$J_i(u_i^*, u_{-i}^*) \geq J_i(u_i, u_{-i}^*), \quad \forall u_i \in U_i, \quad i \in \{1, \ldots, N\},$$  \hfill (2.8)

where $J_i$ is the payoff function of player $i$, $u_i$ its action, $U_i$ its action set, and $u_{-i}$ denotes the actions of the other players. Hence, no player has an incentive to unilaterally deviate its action from $u^*$. In the duopoly example, $U_1 = U_2 = \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the set of positive real numbers.

To attain the Nash strategies (2.6)–(2.7) without any knowledge of modeling information, such as the consumer’s preference $p$, the total demand $S_d$, or the other firm’s marginal cost or price, the firms implement a non-model based real-time optimization strategy, e.g., deterministic extremum seeking with sinusoidal perturbations, to set their price levels. Specifically, P1 and P2 set their prices, $u_1$ and $u_2$ respectively, according to the time-varying strategy (Fig. 2.1):

$$\dot{u}_i(t) = k_i \mu_i(t) J_i(t),$$  \hfill (2.9) \\
$$u_i(t) = \hat{u}_i(t) + \mu_i(t),$$  \hfill (2.10)
where \( \mu_i(t) = a_i \sin(\omega_i t + \varphi_i) \), \( k_i, a_i, \omega_i > 0 \), and \( i \in \{1, 2\} \). Further, the frequencies are of the form,

\[
\omega_i = \omega \bar{\omega}_i, \quad (2.11)
\]

where \( \omega \) is a positive, real number and \( \bar{\omega}_i \) is a positive, rational number. This form is convenient for the convergence analysis performed in Section 2.2. The resulting pricing, sales, and profit transients when the players implement (2.9)–(2.10) are shown in Fig. 2.2 for a simulation with \( S_d = 100 \), \( p = 0.2 \), \( m_1 = m_2 = 30 \), \( a_1 = 0.075 \), \( a_2 = 0.05 \), \( k_1 = 2 \), \( k_2 = 5 \), \( \omega_1 = 26.75 \text{ rad/s} \), \( \omega_2 = 22 \text{ rad/s} \), and \( u_1(0) = \dot{u}_1(0) = 50 \), \( u_2(0) = \dot{u}_2(0) = u_2^* = 110/3. \) From Fig. 2.2(a), we see that convergence to the Nash equilibrium is not trivially achieved since \( u_2(0) = u_2^* \) and yet \( u_2(t) \) increases initially before decreasing to \( u_2^* \) due to the overall system dynamics of (2.9)–(2.10) with (2.4), (2.5). In these simulations (and those shown in Section 2.3.2), we utilize a washout (high pass) filter in each player’s extremum seeking loop [3], but we do not include this filter in the theoretical derivations since it is not needed to derive our stability results. Its inclusion would substantially lengthen the presentation due to the doubling of the state dimension and obfuscate the stability results. The washout filter removes the DC component from signal \( J_i(t) \), which, while not necessary, typically improves performance.

In contrast, the firms are also guaranteed to converge to the Nash equilib-
Figure 2.2: The (a) price, (b) sales, and (c) profit time histories for P1 and P2 when implementing the Nash seeking scheme (2.9)–(2.10). The dashed lines indicate the values at the Nash equilibrium.

which requires each firm to know both its own marginal cost and the other firm’s price at the previous step of the iteration, and also requires P1 to know the total demand $S_d$ and the consumer preference parameter $p$. In essence, P1 must know nearly all the relevant modeling information. When using the extremum
seeking algorithm (2.9)–(2.10), the firms only need to measure the value of their own payoff functions, $J_1$ and $J_2$. Convergence of (2.12), (2.13) is global, whereas the convergence of the Nash seeking strategy for this example can be proved to be semi-global, following [103], or locally, by applying the theory of averaging [55]. We establish local results since we extend this work to include non-quadratic payoff functions in Chapter 3. We do, however, state a non-local result for static games with quadratic payoff functions using the theory found in [106]. For detailed analysis of the non-local convergence of extremum seeking controllers applied to general convex systems, the reader is referred to [103].

2.2 $N$-Player Games with Quadratic Payoff Functions

We now generalize the duopoly example in Section 2.1 to static noncooperative games with $N$ players that wish to maximize their quadratic payoff functions. We prove convergence to a neighborhood of the Nash equilibrium when the players employ the Nash seeking strategy (2.9)–(2.10).

2.2.1 General Quadratic Games

First, we consider games with general quadratic payoff functions. Specifically, the payoff function of player $i$ is of the form,

$$J_i(t) = \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i u_j(t) u_k(t) + \sum_{j=1}^{N} d_j^i u_j(t) + c_i^i,$$

(2.14)

where the action of player $j$ is $u_j \in U_j = \mathbb{R}$, $D_{jk}^i$, $d_j^i$, and $c_i^i$ are constants, $D_{ii}^i < 0$, and $D_{jk}^i = D_{kj}^i$. Quadratic games of this form are studied in Section 4.6 in [11] where Proposition 4.6 states that the $N$-player game with payoff functions (2.14) admits a Nash equilibrium $u^* = [u_1^*, \ldots, u_N^*]^T$ if and only if

$$D_{ii}^i u_i^* + \sum_{j \neq i} D_{ij}^i u_j^* + d_i^i = 0, \quad i \in \{1, \ldots, N\},$$

(2.15)
admits a solution. Rewritten in matrix form, we have \( Du^* = -d \), where

\[
D \triangleq \begin{bmatrix}
D_{11}^1 & D_{12}^1 & \cdots & D_{1N}^1 \\
D_{21}^2 & D_{22}^2 & \cdots & \\
\vdots & \ddots & \ddots & \\
D_{N1}^N & D_{N2}^N & \cdots & D_{NN}^N
\end{bmatrix}, \quad d \triangleq \begin{bmatrix}
d_1^1 \\
d_2^2 \\
\vdots \\
d_N^N
\end{bmatrix},
\tag{2.16}
\]

and \( u^* \) is unique if \( D \) is invertible. We make the following stronger assumption concerning this matrix:

**Assumption 2.1** The matrix \( D \) given by (2.16) is strictly diagonally dominant, i.e.,

\[
\sum_{j \neq i}^N |D_{ij}^i| < |D_{ii}^i|, \quad i \in \{1, \ldots, N\}.
\tag{2.17}
\]

By Assumption 2.1, the Nash equilibrium \( u^* \) exists and is unique since strictly diagonally dominant matrices are nonsingular by the Levy-Desplanques theorem [48, 105]. To attain \( u^* \) stably in real time, without any modeling information, each player \( i \) employs the extremum seeking strategy (2.9)–(2.10).

**Theorem 2.1** Consider the system (2.9)–(2.10) with (2.14) under Assumption 2.1 for an \( N \)-player game, where \( \omega_i \neq \omega_j, 2\omega_i \neq \omega_j, \) and \( \omega_i \neq \omega_j + \omega_k \) for all distinct \( i, j, k \in \{1, \ldots, N\} \), and where \( \omega_i/\omega_j \) is rational for all \( i, j \in \{1, \ldots, N\} \). There exist \( \omega^*, M, m > 0 \) such that for all \( \omega > \omega^* \), if \( |\Delta(0)| \) is sufficiently small, then for all \( t \geq 0 \),

\[
|\Delta(t)| \leq Me^{-mt}|\Delta(0)| + O \left( \frac{1}{\omega} + \max_i a_i \right),
\tag{2.18}
\]

where \( \Delta(t) = [u_1(t) - u^*_1, \ldots, u_N(t) - u^*_N]^T \) and \(|\cdot|\) denotes the Euclidean norm.

**Proof:** Denote the relative Nash equilibrium error as

\[
\tilde{u}_i(t) = u_i(t) - \mu_i(t) - u^*_i,
\]

\[
= \hat{u}_i(t) - u^*_i.
\tag{2.19}
\]
By substituting (2.14) into (2.9)–(2.10), we get the error system,

\[
\dot{\tilde{u}}_i(t) = k_i \mu_i(t) \left( \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\tilde{u}_j(t) + u_j^* + \mu_j(t)) (\tilde{u}_k(t) + u_k^* + \mu_k(t)) \right.
\]

\[
+ \sum_{j=1}^{N} d_j^i (\tilde{u}_j(t) + u_j^* + \mu_j(t)) + c^i \Bigg) .
\]

(2.20)

Let \( \tau = \omega t \) where \( \omega \) is the positive, real number in (2.11). Rewriting (2.20) in the time scale \( \tau \) and rearranging terms yields

\[
\frac{d\tilde{u}_i(\tau)}{d\tau} = \frac{k_i \mu_i(\tau)}{2\omega} \left[ \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\tilde{u}_j(\tau) + u_j^*) (\tilde{u}_k(\tau) + u_k^*) \right.
\]

\[
+ 2 \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\tilde{u}_j(\tau) + u_j^*) \mu_k(\tau) + \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i \mu_j(\tau) \mu_k(\tau)
\]

\[
+ 2 \sum_{j=1}^{N} d_j^i (\tilde{u}_j(\tau) + u_j^* + \mu_j(\tau)) + 2c^i \Bigg],
\]

(2.21)

where \( \mu_i(\tau) = a_i \sin(\bar{\omega}_i \tau + \varphi_i) \) and \( \bar{\omega}_i \) is a rational number. Hence, the error system (2.21) is periodic with period \( T = 2\pi \times \text{LCM} \{1/\bar{\omega}_1, \ldots, 1/\bar{\omega}_N\} \), where LCM denotes the least common multiple. With \( 1/\omega \) as a small parameter, (2.21) admits the application of the averaging theory [55] for stability analysis. The average error system can be shown to be

\[
\frac{d}{d\tau} \tilde{u}^{\text{ave}}_i = \frac{1}{\omega} \left( \frac{1}{T} \int_0^T f_i(\tau, \tilde{u}^{\text{ave}}_1, \ldots, \tilde{u}^{\text{ave}}_N, 0) \, d\tau \right),
\]

\[
= \frac{1}{2\omega} k_i a_i^2 \sum_{j=1}^{N} D_{ij}^i \tilde{u}^{\text{ave}}_j(\tau),
\]

(2.22)

which in matrix form is \( d\tilde{u}^{\text{ave}}/d\tau = A\tilde{u}^{\text{ave}} \), where

\[
A = \frac{1}{2\omega} \begin{bmatrix}
\kappa_1 D_{11}^1 & \kappa_1 D_{12}^1 & \cdots & \kappa_1 D_{1N}^1 \\
\kappa_2 D_{21}^1 & \kappa_2 D_{22}^1 & \cdots & \kappa_2 D_{2N}^1 \\
\vdots & \vdots & \ddots & \vdots \\
\kappa_N D_{N1}^1 & \kappa_N D_{N2}^1 & \cdots & \kappa_N D_{NN}^1
\end{bmatrix},
\]

(2.23)
and $\kappa_i = k_i a_i^2$, $i \in \{1, \ldots, N\}$. (The details of computing (2.22), which require that $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$, are shown in Section A.1 of Appendix A.)

From the Gershgorin Circle Theorem [48, Theorem 6.1.1], we have $\lambda(A) \subseteq \bigcup_{i=1}^{N} \rho_i$, where $\lambda(A)$ denotes the spectrum of $A$ and $\rho_i$ is a Gershgorin disc,

$$\rho_i = \frac{k_i a_i^2}{2\omega} \left\{ z \in \mathbb{C} \left| z - D_{ii}^i < \sum_{j \neq i} |D_{ij}| \right. \right\}.$$  

(2.24)

Since $D_{ii}^i < 0$ and $D$ is strictly diagonally dominant, the union of the Gershgorin discs lies strictly in the left half of the complex plane, and we conclude that $\text{Re}\{\lambda\} < 0$ for all $\lambda \in \lambda(A)$. Thus, given any matrix $Q = Q^T > 0$, there exists a matrix $P = P^T > 0$ satisfying the Lyapunov equation $PA + A^T P = -Q$.

Using $V(\tau) = (\tilde{u}^{\text{ave}}(\tau))^T P \tilde{u}^{\text{ave}}(\tau)$ as a Lyapunov function, we obtain

$$\dot{V} = -(\tilde{u}^{\text{ave}})^T Q \tilde{u}^{\text{ave}} \leq -\lambda_{\text{min}}(Q) |\tilde{u}^{\text{ave}}|^2.$$  

(2.25)

Noting that $V$ satisfies the bounds, $\lambda_{\text{min}}(P) |\tilde{u}^{\text{ave}}(\tau)|^2 \leq V(\tau) \leq \lambda_{\text{max}}(P) |\tilde{u}^{\text{ave}}(\tau)|^2$, and applying the Comparison Lemma [55] gives

$$|\tilde{u}^{\text{ave}}(\tau)| \leq M e^{-\frac{m}{\omega}} |\tilde{u}^{\text{ave}}(0)|,$$  

(2.26)

where $M = \sqrt{\frac{\lambda_{\text{max}}(P)}{\lambda_{\text{min}}(P)}}$, $m = \omega \frac{\lambda_{\text{min}}(Q)}{2\lambda_{\text{max}}(P)}$.  

(2.27)

From (2.26) and [55, Theorem 10.4], we obtain $|\tilde{u}(\tau)| \leq M e^{-\frac{m}{\omega}} |\tilde{u}(0)| + O(1/\omega)$, provided $\tilde{u}(0)$ is sufficiently close to $\tilde{u}^{\text{ave}}(0)$. Reverting to the time scale $t$ and noting that $u_i(t) - u_i^* = \bar{u}_i(t) + \mu_i(t) = \bar{u}_i(t) + O(\max_i a_i)$ completes the proof. 

From the proof, we see that the convergence result holds if (2.23) is Hurwitz, which does not require the strict diagonal dominance assumption. However, Assumption 2.1 allows convergence to hold for $k_i, a_i > 0$, whereas merely assuming (2.23) is Hurwitz would create a potentially intricate dependence on the unknown game model and the selection of the parameters $k_i, a_i$.

While we have considered only the case where the action variables of the players are scalars, the results equally apply to the vector case, namely $u_i \in \mathbb{R}^n$, by simply considering each different component of a player’s action variable to be
controlled by a different (virtual) player. In this case, the payoff functions of all virtual players corresponding to player $i$ will be the same.

Even though $u^*$ is unique for quadratic payoffs, Theorem 2.1 is local due to our use of standard local averaging theory. From the theory in [106], we have the following non-local result:

**Corollary 2.1** Consider $N$ players with quadratic payoff functions (2.14) that implement the Nash seeking strategy (2.9)–(2.10) with frequencies satisfying the inequalities stated in Theorem 2.1. Then, the Nash equilibrium $u^*$ is semi-globally practically asymptotically stable.

**Proof:** In the proof of Theorem 2.1, the average error system (2.22) is shown to be globally asymptotically stable. By [106, Theorem 2], with the error system (2.21) satisfying the theorem’s conditions, the origin of (2.21) is semi-globally practically asymptotically stable.

For more details on semi-global convergence with extremum seeking controllers, the reader is referred to [103].

### 2.2.2 Symmetric Quadratic Games

If the matrix $D$ is symmetric, we can develop a more precise expression for the convergence rate in Theorem 2.1. Specifically, we assume the following:

**Assumption 2.2** $D^i_{ij} = D^j_{ji}$ for all $i, j \in \{1, \ldots, N\}$.

Under Assumptions 2.1 and 2.2, $D$ is a negative definite symmetric matrix. Noncooperative games satisfying 2.2 are a class of games known as potential games [73, Theorem 4.5]. In potential games, the maximization of the players’ payoff functions $J_i$ corresponds to the maximization of a global function $\Phi(u_1, \ldots, u_N)$, known as the potential function.

**Corollary 2.2** Consider the system (2.9)–(2.10) with payoff functions (2.14) under Assumptions 2.1 and 2.2 for an $N$-player game, where $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$, and where $\omega_i/\omega_j$ is rational
for all \( i, j \in \{1, \ldots, N\} \). The convergence properties of Theorem 2.1 hold with

\[
M = \sqrt{\frac{\max_i \{k_i a_i^2\}}{\min_i \{k_i a_i^2\}}} \tag{2.28}
\]

\[
m = \frac{1}{2} \min_i \{k_i a_i^2\} \min_i \left\{ -D_{ii}^i - \sum_{j \neq i} |D_{ij}^i| \right\} \tag{2.29}
\]

**Proof:** From the proof of Theorem 2.1, given any matrix \( Q = Q^T > 0 \), there exists a matrix \( P = P^T > 0 \) satisfying the Lyapunov equation \( PA + A^T P = -Q \) since \( A \), given by (2.23), is Hurwitz. Under Assumption 2.2, we select \( Q = -D \) and obtain \( P = \text{diag}(\omega/k_1 a_1^2, \ldots, \omega/k_N a_N^2) \). Then, we directly have

\[
\lambda_{\text{min}}(P) = \frac{\omega}{\max_i \{k_i a_i^2\}}, \quad \lambda_{\text{max}}(P) = \frac{\omega}{\min_i \{k_i a_i^2\}} \tag{2.30}
\]

and using the Gershgorin Circle Theorem [48, Theorem 6.1.1], we obtain the bound

\[
\lambda_{\text{min}}(Q) \geq \min_i \left\{ -D_{ii}^i - \sum_{j \neq i} |D_{ij}^i| \right\} \tag{2.31}
\]

where we note that \( D_{ii}^i < 0 \). From (2.27), (2.30), and (2.31), we obtain the result.

Of note, the coefficient \( M \) in Corollary 2.2 is determined completely by the extremum seeking parameters \( k_i, a_i \) while the convergence rate \( m \) depends on both \( k_i, a_i \), and the unknown game parameters \( D_{ij}^i \). Thus, the convergence rate cannot be fully known since it depends on the unknown structure of the game.

### 2.3 Examples

We now revisit the duopoly price game in Section 2.1 and apply our convergence result before studying an oligopoly price game example with four players.

#### 2.3.1 Duopoly Price Game

Revisiting the duopoly example in Section 2.1, we see that the profit functions (2.4), (2.5) have the form (2.14). Moreover, this game satisfies both Assump-
tions 2.1 and 2.2 since
\[ D = \frac{1}{p} \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}. \]

By Theorem 2.1 and Corollary 2.2, the firms converge to a neighborhood of \((u_1^*, u_2^*)\) (2.6)–(2.7) according to
\[
\left| \begin{bmatrix} u_1(t) - u_1^* \\ u_2(t) - u_2^* \end{bmatrix} \right| \leq M e^{-mt} \left| \begin{bmatrix} u_1(0) - u_1^* \\ u_2(0) - u_2^* \end{bmatrix} \right| + O \left( \frac{1}{\omega} + \max(a_1, a_2) \right),
\]
where \(M\) is given by (2.28) and \(m = \min(k_1 a_1^2, k_2 a_2^2)/2p\) since \(\lambda_{\min}(Q) = 1/p\).

### 2.3.2 Oligopoly Price Game

Consider a static noncooperative game with \(N\) firms in an oligopoly market structure that compete to maximize their profits by setting the price \(u_i\) of their product. As in the duopoly example in Section 2.1, let \(m_i\) be the marginal cost of player \(i\), \(s_i\) its sales volume, and \(J_i\) its profit, given by (2.1). The simplified sales model for the duopoly (2.2)–(2.3), where one firm has a clear advantage over the other in terms of consumer preference is no longer appropriate. Instead, we model the sales volume \(s_i\) as
\[
s_i(t) = \frac{R_i}{R_i + \left( S_d - \frac{u_i(t)}{R_i} + \sum_{j \neq i}^N \frac{u_j(t)}{R_j} \right)}, \tag{2.32}
\]
where \(S_d\) is the total consumer demand, \(R_i > 0\) for all \(i\), \(1/R_\parallel = \left( \sum_{k=1}^N 1/R_k \right)\), and \(1/R_i = \left( \sum_{k \neq i}^N 1/R_k \right)\).

The sales model (2.32) is motivated by an analogous electric circuit, shown in Fig. 2.3, where \(S_d\) is an ideal current generator, \(u_i\) are ideal voltage generators, and most importantly, the resistors \(R_i\) represent the “resistance” that consumers have toward buying product \(i\). This resistance may be due to quality or brand image considerations—the most desirable products have the lowest \(R_i\). The sales in (2.32) are inversely proportional to \(R_i\) and grow as \(u_i\) decreases and as \(u_j, j \neq i\), increases. The profit (2.1), in electrical analogy, corresponds to the power absorbed by the \(u_i - m_i\) portion of the voltage generator \(i\).
Figure 2.3: A model of sales $s_1$, $s_2$, $s_3$ in a three-firm oligopoly with prices $u_1$, $u_2$, $u_3$ and total consumer demand $S_d$. The desirability of product $i$ is proportional to $1/R_i$.

**Proposition 1** There exists $\omega^*$ such that, for all $\omega > \omega^*$, the system (2.9)–(2.10) with (2.1) and (2.32) for the oligopoly price game, where $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$ for all distinct $i, j = \{1, \ldots, N\}$ and where $\omega_i/\omega_j$ is rational for all $i, j \in \{1, \ldots, N\}$, exponentially converges to a neighborhood of the Nash equilibrium,

$$u^*_i = \frac{\prod R_i}{2R_i + R_i} \left( R_i S_d + m_i + \sum_{j=1}^{N} \frac{m_j R_i - m_i R_j}{2R_j + R_j} \right),$$  

(2.33)

where $\Pi^{-1} = 1 - \sum_{j=1}^{N} R_j/(2R_j + R_j) > 0$. Namely, if $|\Delta(0)|$ is sufficiently small, then for all $t \geq 0$,

$$|\Delta(t)| \leq M e^{-mt} |\Delta(0)| + O \left( \frac{1}{\omega} + \max_i (a_i) \right),$$  

(2.34)

where $\Delta$ is defined in Theorem 2.1, $M$ is given by (2.28), and

$$m = \frac{R_{\parallel} \min_i \{k_i a_i^2\}}{2 \max_i \{R_i \Gamma_i\}}, \quad \Gamma_i = \min_{j \in \{1, \ldots, N\}, j \neq i} R_j,$$  

(2.35)

**Proof:** Substituting (2.32) into (2.1) yields payoff functions,

$$J_i(t) = \frac{R_{\parallel}}{R_i} \left( -\frac{u_i^2}{R_i} + u_i \sum_{j \neq i} \frac{u_j}{R_j} + \left( \frac{m_i}{R_i} + S_d \right) u_i - m_i \sum_{j \neq i} \frac{u_j}{R_j} - S_d m_i \right),$$

that are of the form (2.14). Therefore, the Nash equilibrium $u^*$ satisfies $Du^* = -d$, where $D$ and $d$ are given by (2.16) and have elements

$$D_{ij}^i = \begin{cases} -\frac{2R_i}{R_{\parallel} R_i} & \text{if } i = j; \\ \frac{R_i}{R_{\parallel} R_j} & \text{if } i \neq j; \end{cases} \quad d_i^j = \frac{m_j R_{\parallel}}{R_j R_j} + \frac{S R_i}{R_i R_j},$$
for \(i, j \in \{1, \ldots, N\}\). Assumption 2.1 is satisfied by \(D\) since
\[
\sum_{j \neq i}^N \frac{R_{ij}}{|R_i - R_j|} = \frac{R_i}{R_i - R_j} < \frac{-2R_i}{|R_i R_j|}, \quad i \in \{1, \ldots, N\}.
\]
Thus, the Nash equilibrium of this game exists, is unique, and can be shown to be (2.33). (The various parameters here are assumed to be selected such that \(u_i^*\) is positive for all \(i\).) Moreover, \(D_{ij}^i = D_{ji}^j\), so \(D\) is a negative definite symmetric matrix, satisfying Assumption 2.2. Then, by Theorem 2.1 and Corollary 2.2, we have the convergence bound (2.34) and obtain \(m\) by computing
\[
\lambda_{\min}(Q) \geq \min_i \left\{ \frac{2R_i}{R_i R_i} - \sum_{j \neq i} \frac{R_i R_j}{R_i R_j} \right\} = \frac{R_i}{\max_i \{R_i R_i\}},
\]
where \(Q = -D\), and by noting that \(\max_i \{R_i R_i\} < \max_i \{R_i \Gamma_i\}\). Finally, the error system (2.20) for this game does not contain any terms with the product \(\mu_i(t)\mu_j(t)\mu_k(t)\), so the requirement that \(\omega_i \neq \omega_j + \omega_k\) for all distinct \(i, j, k \in \{1, \ldots, N\}\) does not arise when computing the average error system.

The resulting pricing, sales, and profit transients when four firms implement (2.9)–(2.10) are shown in Fig. 2.4 for a simulation with game parameters: \(S_d = 100, R_1 = 0.15, R_2 = 0.3, R_3 = .6, R_4 = 1, m_1 = 30, m_2 = 30, m_3 = 25, m_4 = 20\); extremum seeking parameters: \(a_1 = a_2 = a_3 = a_4 = 0.05, k_1 = 6, k_2 = 18, k_3 = 10, k_4 = 24, \omega_1 = 30, \omega_2 = 24, \omega_3 = 44, \omega_4 = 36\); and initial conditions: \(u_1(0) = \hat{u}_1(0) = 52, u_2(0) = \hat{u}_2(0) = u_2^* = 40.93, u_3(0) = \hat{u}_3(0) = 33.5, u_4(0) = \hat{u}_4(0) = u_4^* = 35.09\).

### 2.4 Approximate Best Response Strategy

We have established that the Nash seeking scheme (2.9)–(2.10) converges to a Nash equilibrium, but we have not studied the transient response of the players’ when utilizing this learning strategy. Further investigation shows that this scheme approximates, on the average, a best response strategy with a player’s best response dynamics scaled by a parameter of the unknown payoff function. To remove this scaling, we introduce the Newton-based Nash seeking strategy.
Figure 2.4: The (a) price, (b) sales, and (c) profit time histories of firms P1, P2, P3, and P4 when implementing the Nash seeking strategy (2.9)–(2.10). The dashed lines indicate the values at the Nash equilibrium.

2.4.1 N-Player Games with Stubborn Players

An interesting scenario to consider is when not all the players utilize the Nash seeking strategy (2.9)–(2.10). Without loss of generality, we assume that players $i \in \{1, \ldots, n\}$ implement (2.9)–(2.10) while players $j \in \{n+1, \ldots, N\}$ are stubborn and use fixed actions,

$$u_j(t) \equiv \bar{u}_j, \quad j \in \{n+1, \ldots, N\}. \quad (2.36)$$

To discuss the convergence of the Nash seeking players, we introduce their reaction curves [11, Definition 4.3], which are the players’ best response strategy
given the actions of the other players. When the players have quadratic payoff functions (2.14), their reaction curves are

\[ l_i(u_{-i}) = -\frac{1}{D_{ii}} \left( \sum_{j \neq i}^N D_{ij} u_j + d_i \right), \]

(2.37)

where \( u_{-i} \) denotes the actions of the other \( N - 1 \) players. In the presence of stubborn players, the remaining players’ unique best response \( u^{br} = [u_1^{br}, \ldots, u_n^{br}]^T \) is given by

\[
\begin{bmatrix}
  u_1^{br} \\
  \vdots \\
  u_n^{br}
\end{bmatrix} = - \begin{bmatrix} D_{11} & \cdots & D_{1n} \\
  \vdots & \ddots & \vdots \\
  D_{n1} & \cdots & D_{nn}
\end{bmatrix}^{-1} \left( \begin{bmatrix} D_{1,n+1} & \cdots & D_{1,N} \\
  \vdots & \ddots & \vdots \\
  D_{n,n+1} & \cdots & D_{n,N}
\end{bmatrix} \begin{bmatrix} \bar{u}_{n+1} \\
  \vdots \\
  \bar{u}_N
\end{bmatrix} + \begin{bmatrix} d_1^n \\
  \vdots \\
  d_n^n
\end{bmatrix} \right).
\]

(2.38)

When there are no stubborn players, \( u^{br} = u^* \).

**Theorem 2.2** Consider the \( N \)-player game with (2.14) under Assumption 2.1, where players \( i \in \{1, \ldots, n\} \) implement the strategy (2.9)–(2.10), with \( \omega_i \neq \omega_j, 2\omega_i \neq \omega_j, \) and \( \omega_i \neq \omega_j + \omega_k \) for all distinct \( i, j, k \in \{1, \ldots, n\} \), and stubborn players \( j \in \{n + 1, \ldots, N\} \) implement (2.36). There exist \( \omega^*, M, m > 0 \) such that for all \( \omega > \omega^* \), if \( |\Omega(0)| \) is sufficiently small, then for all \( t \geq 0 \),

\[ |\Omega(t)| \leq Me^{-mt}|\Omega(0)| + O \left( \frac{1}{\omega} + \max_i a_i \right), \]

(2.39)

where \( \Omega(t) = [u_1(t) - u_1^{br}, \ldots, u_n(t) - u_n^{br}]^T \) and \( u^{br} \) is given by (2.38).

**Proof:** Following the proof of Theorem 2.1 for the \( n \) Nash seeking players, one can obtain

\[
\frac{d}{d\tau} \begin{bmatrix} \bar{u}_1^{ave} \\
  \vdots \\
  \bar{u}_n^{ave}
\end{bmatrix} = \frac{1}{2\omega} \begin{bmatrix} \kappa_1 D_{11} & \cdots & \kappa_1 D_{1n} \\
  \vdots & \ddots & \vdots \\
  \kappa_n D_{n1} & \cdots & \kappa_n D_{nn}
\end{bmatrix} \begin{bmatrix} \bar{u}_1^{ave} \\
  \vdots \\
  \bar{u}_n^{ave}
\end{bmatrix} + \frac{1}{2\omega} \begin{bmatrix} \kappa_1 D_{1,n+1} & \cdots & \kappa_1 D_{1,N} \\
  \vdots & \ddots & \vdots \\
  \kappa_n D_{n,n+1} & \cdots & \kappa_n D_{n,N}
\end{bmatrix} \begin{bmatrix} \bar{v}_{n+1} \\
  \vdots \\
  \bar{v}_N
\end{bmatrix},
\]

(2.40)
where \( \kappa_i = k_i a_i^2 \), \( i \in \{1, \ldots, n\} \) and the stubborn players’ error relative to the Nash equilibrium is denoted by \( \tilde{v}_j = \bar{u}_j - u_j^* \), \( j \in \{n+1, \ldots, N\} \). The unique, exponentially stable equilibrium of (2.40) is

\[
\begin{bmatrix}
\tilde{u}_1^e \\
\vdots \\
\tilde{u}_n^e
\end{bmatrix} = - \begin{bmatrix} D_{11} & \cdots & D_{1n} \\
\vdots & \ddots & \vdots \\
D_{n1} & \cdots & D_{nn} \end{bmatrix}^{-1} \begin{bmatrix} D_{1,n+1} & \cdots & D_{1N} \\
\vdots & \ddots & \vdots \\
D_{n,n+1} & \cdots & D_{nN} \end{bmatrix} \begin{bmatrix} \tilde{v}_{n+1} \\
\vdots \\
\tilde{v}_N \end{bmatrix},
\]

(2.41)

and so from the proof of Theorem 2.1 and [55, Theorem 10.4], we have

\[
|u(t) - \tilde{u}^e - u^*| \leq Me^{-mt}|u(0) - \tilde{u}^e - u^*| + O \left( \frac{1}{\omega} + \max_i a_i \right),
\]

(2.42)

for players \( i \in \{1, \ldots, n\} \). What remains to be shown is that \( \tilde{u}_i^e + u_i^* \) is the best response of player \( i \).

From (2.36) and (2.42), the players’ actions converge on average to \( u_{ss} = [\bar{u}_1^e + u_1^*, \ldots, \bar{u}_n^e + u_n^*, \bar{u}_{n+1}, \ldots, \bar{u}_N]^T \). At \( u_{ss} \), the best response of player \( i \in \{1, \ldots, n\} \) is

\[
l_i(u_{ss}) = -\frac{1}{D_{ii}} \left( \sum_{j \neq i}^n D_{ij}^i (\bar{u}_j^e + u_j^*) + \sum_{j=n+1}^N D_{ij}^i \bar{u}_j + d_i^j \right),
\]

\[
= -\frac{1}{D_{ii}} \left( \sum_{j \neq i}^n D_{ij}^i \tilde{u}_j^* + \sum_{j=n+1}^N D_{ij}^i \bar{v}_j \right) - \frac{1}{D_{ii}} \left( \sum_{j \neq i}^N D_{ij}^i u_j^* + d_i^j \right),
\]

(2.43)

where we have substituted \( \bar{u}_j = \bar{v}_j + u_j^* \). Noting (2.37) and using (2.41) to substitute for \( \bar{v}_j \) yields

\[
l_i(\tilde{u}_{ss}) = -\frac{1}{D_{ii}} \left( \sum_{j \neq i}^n D_{ij}^i \tilde{u}_j^e - \sum_{j=1}^n D_{ij}^i \bar{u}_j^e \right) + u_i^*
\]

\[
= \tilde{u}_i^e + u_i^*.
\]

(2.44)

Hence, the \( n \) Nash seeking players converge to their best response actions \( u_{br} \).

To provide more insight into this result, we note that the average response
of player $i$ can be shown to be

$$
\dot{\bar{u}}_i^{\text{ave}}(t) = \frac{k_i a_i^2}{2} \frac{\partial J_i}{\partial u_i} (\hat{\bar{u}}^{\text{ave}}(t)),
$$

$$
= \frac{k_i a_i^2}{2} \left( \sum_{j=1}^{N} D_{ij} \hat{\bar{u}}^{\text{ave}}_j(t) + d_i^i \right),
$$

$$
= \frac{k_i a_i^2}{2} |D_{ii}| \left( l_i (\bar{u}_i^{\text{ave}}(t) - \hat{\bar{u}}^{\text{ave}}_i(t)) \right),
$$

(2.45)

which is the scaled, continuous-time best response for games with quadratic payoff functions. Continuous-time best response dynamics are studied for more general scenarios in [46, 47].

Revisiting the oligopoly price game in Section 2.3.2, Figure 2.5 depicts the time histories of the firms’ prices, sales, and profits when firm P1 is a stubborn player, whose price is given by,

$$
\bar{u}_1 = \begin{cases} 
50 & \text{if } 0 < t \leq 75, \\
46 & \text{if } 75 < t \leq 150, \\
u_1^* = 42.88 & \text{if } 150 < t \leq 200.
\end{cases}
$$

(2.46)

and firms P2, P3, and P4 are Nash seeking players who use the same parameter values listed in Section 2.3.2.

The time evolution of the firms’ profits (Figure 2.5(c)) shows that the stubborn firm P1 benefits by discounting its good from $\bar{u}_1 = 50$ to $\bar{u}_1 = 46$, but not from $\bar{u}_1 = 46$ to $\bar{u}_1 = u_1^*$. Hence, a stubborn player in this scenario could take on a leadership role by setting its action, assuming the other players invoke their best responses, and then, after the Nash seeking players’ actions reach a steady state value (which the leader can estimate from its measured profit values), setting a new constant action in an attempt to converge iteratively to a maximum profit value. This two-stage structure is akin to a Stackelberg game played over the firms’ steady state values.

### 2.4.2 Newton-Based Nash Seeking Strategy

Of note in (2.45) is that the convergence rate of player $i$’s response depends on the potentially unknown, and likely uncertain, value $|D_{ii}|$. This dependence
on an unknown value, however, can be removed by utilizing a more sophisticated, Newton-based Nash seeking strategy that estimates the second-order partial derivative $\partial^2 J_i / \partial u_i^2$.

The Newton-based Nash seeking strategy is given by

\[
\dot{\hat{u}}_i(t) = -\frac{2k_i}{\sigma_i^2} \gamma_i(t) \mu_i(t) J_i(t),
\]

\[
\dot{\gamma}_i(t) = \sigma_i \gamma_i(t) - \sigma_i \gamma_i^2(t) \nu_i(t) J_i(t),
\]

\[
u_i(t) = \dot{\hat{u}}_i(t) + \mu_i(t),
\]
where \( \mu_i(t) = a_i \sin(\omega_i t + \varphi_i), \nu_i(t) = (16/a_i^2)(\sin^2(\omega_i t + \varphi_i) - 1/2), \) and \( k_i, a_i, \sigma_i, \omega_i > 0 \) for \( i \in \{1, \ldots, N\} \). As in (2.11), the frequencies are of the form \( \omega_i = \omega \bar{\omega}_i \), where \( \omega \) is a positive, real number and \( \bar{\omega}_i \) is a positive, rational number. For vector-valued actions, \( \gamma_i(t) \) and \( \nu_i(t) \) become matrix-valued expressions as shown in [42].

**Theorem 2.3** Consider the system (2.47)–(2.49) with (2.14) under Assumption 2.1 for an \( N \)-player game, where \( \omega_i \neq \omega_j, \omega_i \neq \omega_j + \omega_k, 2\omega_i \neq \omega_j + \omega_k, \) and \( \omega_i \neq 2\omega_j + \omega_k \) for all distinct \( i, j, k \in \{1, \ldots, N\} \), and where \( \omega_i/\omega_j \) is rational for all \( i, j \in \{1, \ldots, N\} \). There exist \( \omega^*, M, m > 0 \) such that for all \( \omega > \omega^* \), if \( ||u(0) - u^*, \gamma(0) - \gamma^*||^T \) is sufficiently small, then for all \( t \geq 0 \),

\[
\begin{bmatrix}
\left| u(t) - u^* \right| \\
\left| \gamma(t) - \gamma^* \right|
\end{bmatrix} \leq Me^{-mt} \begin{bmatrix} \left| u(0) - u^* \right| \\
\left| \gamma(0) - \gamma^* \right|
\end{bmatrix} + O \left( \frac{1}{\omega} + \max_{i} a_i \right),
\]

(2.50)

where \( u(t) = [u_1(t), \ldots, u_N(t)]^T, \ u^* = [u^*_1, \ldots, u^*_N]^T, \ \gamma(t) = [\gamma_1(t), \ldots, \gamma_N(t)]^T, \) and \( \gamma^* = [1/D_{11}, \ldots, 1/D_{NN}]^T \).

**Proof:** Denote the error variables \( \bar{u}_i, \bar{\gamma}_i \) as

\[
\bar{u}_i(t) = u_i(t) - \mu_i(t) - u_i^*, \quad (2.51)
\]

\[
\bar{\gamma}_i(t) = \bar{\gamma}_i(t) - \frac{1}{D_{ii}}, \quad (2.52)
\]

From (2.14), (2.47)–(2.49), (2.51), and (2.52), we obtain, after rearranging terms, the following error system in the time scale \( \tau = \omega t \):

\[
\frac{d}{d\tau} \bar{u}_i(\tau) = -\frac{k_i}{\omega a_i^2} \left( \bar{\gamma}_i(\tau) + \frac{1}{D_{ii}} \right) \mu_i(\tau) \left[ \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\bar{u}_j(\tau) + u_j^*) (\bar{u}_k(\tau) + u_k^*) \right.
\]

\[
+ 2 \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\bar{u}_j(\tau) + u_j^*) \mu_k(\tau) + \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i \nu_j(\tau) \mu_k(\tau)
\]

\[
\left. + 2 \sum_{j=1}^{N} d_{ji}^i (\bar{u}_j(\tau) + u_j^* + \mu_j(\tau)) + 2c \right],
\]

(2.53)

\[
\frac{d}{d\tau} \bar{\gamma}_i(\tau) = \frac{\sigma_i}{\omega} \left( \bar{\gamma}_i(\tau) + \frac{1}{D_{ii}} \right) - \frac{\sigma_i}{2\omega} \left( \bar{\gamma}_i(\tau) + \frac{1}{D_{ii}} \right) \nu_i(t)
\]

\[
\times \left[ \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\bar{u}_j(\tau) + u_j^*) (\bar{u}_k(\tau) + u_k^*) \right].
\]
\[ + 2 \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i (\tilde{u}_j(\tau) + u_j^\ast) \mu_k(\tau) + \sum_{j=1}^{N} \sum_{k=1}^{N} D_{jk}^i \mu_j(\tau) \mu_k(\tau) \]
\[ + 2 \sum_{j=1}^{N} d_j^i (\tilde{u}_j(\tau) + u_j^\ast + \mu_j(\tau)) + 2c^i \], (2.54)

where \( \mu_i(\tau) = a_i \sin(\bar{\omega}_i \tau + \varphi_i) \) and \( \nu_i(\tau) = (16/a_i^2)(\sin^2(\bar{\omega}_i \tau + \varphi_i) - 1/2) \). The error system (2.21) is periodic with period \( T = 2\pi \times \text{LCM} \{1/\bar{\omega}_1, \ldots, 1/\bar{\omega}_N \} \), where LCM denotes the least common multiple, with \( 1/\omega \) as a small parameter, which allows the application of the averaging theory [55] for stability analysis.

The average error system can be shown to be
\[
\frac{d}{d\tau} \tilde{u}_{i\text{ave}} = -\frac{k_i}{\omega} \left( \tilde{\gamma}_{i\text{ave}} + \frac{1}{D_{ii}^i} \right) \sum_{j=1}^{N} D_{ij}^i \tilde{u}_j \text{ave}, \quad (2.55)
\]
\[
\frac{d}{d\tau} \tilde{\gamma}_{i\text{ave}} = \frac{\sigma_i}{\omega} \left( \tilde{\gamma}_{i\text{ave}} + \frac{1}{D_{ii}^i} \right) - \frac{\sigma_i}{\omega} \left( \tilde{\gamma}_{i\text{ave}} + \frac{1}{D_{ii}^i} \right)^2 D_{ii}^i, \quad (2.56)
\]

As in the proof of Theorem 2.1, the details of computing (2.55) are shown in Section A.1 of Appendix A, whereas computing (2.56) requires the integrals shown in Sections A.1 and A.2. The nonlinear ODE (2.56) has two equilibria, \( \tilde{\gamma}_i^e = 0 \) and \( \tilde{\gamma}_i^e = -1/D_{ii}^i \). At \( \tilde{\gamma}_i^e = 0 \), the linearized \( \tilde{\gamma}_{i\text{ave}} \)-system is
\[
\frac{d}{d\tau} (\delta \tilde{\gamma})_{i\text{ave}} = \left[ \frac{\sigma_i}{\omega} - 2 \frac{\sigma_i}{\omega} \left( \frac{1}{D_{ii}^i} \right) D_{ii}^i \right] (\delta \tilde{\gamma})_{i\text{ave}}, \quad (2.57)
\]
which is exponentially stable, and at \( \tilde{\gamma}_i^e = -1/D_{ii}^i \), it is
\[
\frac{d}{d\tau} (\delta \tilde{\gamma})_{i\text{ave}} = \left[ \frac{\sigma_i}{\omega} - 2 \frac{\sigma_i}{\omega} \left( -\frac{1}{D_{ii}^i} + \frac{1}{D_{ii}^i} \right) D_{ii}^i \right] (\delta \tilde{\gamma})_{i\text{ave}}, \quad (2.58)
\]
which is unstable. (Hence, \( 1/D_{ii}^i \) is a stable equilibrium of the \( \gamma_{i\text{ave}} \)-system, and the origin is unstable, which means \( \gamma(0) \) should be initialized less than zero.) We study the stability of the total average system (2.55)–(2.56) by linearizing about the equilibrium \( (\tilde{u}^e, \tilde{\gamma}^e) = (0, 0) \) to obtain, in matrix form,
\[
\frac{d}{d\tau} \begin{bmatrix} \delta \tilde{u}_{\text{ave}} \\ \delta \tilde{\gamma}_{\text{ave}} \end{bmatrix} = \frac{1}{\omega} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} \delta \tilde{u}_{\text{ave}} \\ \delta \tilde{\gamma}_{\text{ave}} \end{bmatrix}, \quad (2.59)
\]
where \( \delta \tilde{u}^{\text{ave}} = [(\delta \tilde{u})^{\text{ave}}_1, \ldots, (\delta \tilde{u})^{\text{ave}}_N]^T, \delta \gamma^{\text{ave}} = [(\delta \gamma)^{\text{ave}}_1, \ldots, (\delta \gamma)^{\text{ave}}_N]^T \),

\[
A_1 = \begin{bmatrix}
-k_1 & -k_1 \frac{D^1_{12}}{D^1_{11}} & \cdots & -k_1 \frac{D^1_{1N}}{D^1_{11}} \\
-k_1 \frac{D^2_{11}}{D^2_{22}} & -k_2 & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
-k_N \frac{D^N_{N1}}{D^N_{NN}} & \cdots & -k_N 
\end{bmatrix},
\]

(2.60)

\[A_2 = \text{diag}[-\sigma_1, \ldots, -\sigma_N],\] and \( \text{diag}[:] \) denotes a diagonal matrix.

For \((\tilde{u}^e, \tilde{\gamma}^e) = (0, 0)\) to be a locally stable equilibrium point of the average system (2.55)–(2.56), \( A_1 \) and \( A_2 \) must be Hurwitz. Since \( A_2 \) is diagonal with negative entries, it is clearly Hurwitz. For \( A_1 \), we rewrite it as the matrix product

\[A_1 = KD,\]

(2.61)

where \( K = \text{diag}[-k_1/D^1_{11}, \ldots, -k_N/D^N_{NN}], D^i_{ii} < 0, \) and \( D \) is the diagonally dominant matrix given by (2.16). Consequently, following the proof of Theorem 2.1 and the Gershgorin Circle Theorem [48, Theorem 6.1.1], \( A_1 \) is Hurwitz with eigenvalues contained in the set \( \bigcup_{i=1}^N \rho_i \), where

\[\rho_i = \frac{k_i}{|D^i_{ii}|} \left\{ z \in \mathbb{C} \mid |z - D^i_{ii}| < \sum_{j \neq i} |D^i_{ij}| \right\} .\]

(2.62)

The remainder of the proof follows from the proof of Theorem 2.1.

When compared to the standard Nash equilibrium seeking scheme, which is shown in (2.45) to be gradient-based, (2.47) and (2.49) are qualitatively the same as (2.9) and (2.10) except for the presence of \( \gamma_i(t) \) in (2.47). The additional constant scaling terms are included to remove the scaling factor \( a_i^2/2 \) from the average response of player \( i \) (cf. (2.45)). The main difference between the two schemes lies in (2.48), where, on the average, the perturbation \( \nu_i(t) \) generates an estimate of \( (\partial^2 J_i/\partial u^2_i)(u) \) and the Riccati equation generates the inverse of this value, which, for players with quadratic payoff functions (2.14), leads to the average response,

\[\dot{\hat{u}}^{\text{ave}}_i(t) = -k_i \left( \frac{\partial^2 J_i}{\partial u^2_i}(\hat{\tilde{u}}^{\text{ave}}_i(t)) \right)^{-1} \frac{\partial J_i}{\partial u_i}(\hat{\tilde{u}}^{\text{ave}}_i(t)),\]
\[ \begin{align*}
= & - \frac{k_i}{D_{ii}} \left( \sum_{j=1}^{N} D_{ij} \hat{u}_{j}^{\text{ave}}(t) + d_i^i \right), \\
= & k_i \left( l_i(\hat{u}_{i}^{\text{ave}}(t) - \hat{u}_{i}^{\text{ave}}(t)) \right),
\end{align*} \]
which is scaled only by the known quantity \( k_i \).

### 2.4.3 \( N \)-Player Games with Gradient-Based and Newton-Based Nash Seeking Players

We now consider \( N \)-player games with both gradient-based and Newton-based Nash seeking players with quadratic payoff functions (2.14) to determine if these two non-model based strategies conflict with one another. Without loss of generality, we assume that players \( i \in \{1, \ldots, n\} \) employ gradient-based learning (2.9)–(2.10) but with (2.9) rescaled to remove the factor \( a_i^2/2 \) from the average system, i.e., the players implement
\[ \begin{align*}
\dot{\hat{u}}_i(t) &= \frac{2k_i}{a_i} \mu_i(t) J_i(t), \\
u_i(t) &= \hat{u}_i(t) + \mu_i(t).
\end{align*} \]
Players \( j \in \{n+1, \ldots, N\} \) utilize the Newton-based learning strategy (2.47)–(2.49).

**Corollary 2.3** Consider the \( N \)-player game with (2.14) under Assumption 2.1 where players \( i \in \{1, \ldots, n\} \) implement the strategy (2.64)–(2.65) and players \( j \in \{n+1, \ldots, N\} \) implement (2.47)–(2.49), where \( \omega_i \neq \omega_j, \omega_i \neq \omega_j + \omega_k, 2\omega_i \neq \omega_j + \omega_k, \) and \( \omega_i \neq 2\omega_j + \omega_k \) for all distinct \( i, j, k \in \{1, \ldots, N\} \), and where \( \omega_i/\omega_j \) is rational for all \( i, j \in \{1, \ldots, N\} \). There exist \( \omega^*, M, m > 0 \) such that for all \( \omega > \omega^* \), if
\[ \left| [u(0) - u^*, \gamma(0) - \gamma^*]^T \right| \text{ is sufficiently small}, \] then for all \( t \geq 0 \),
\[ \begin{align*}
\begin{bmatrix} u(t) - u^* \\ \gamma(t) - \gamma^* \end{bmatrix} \leq M e^{-mt} \begin{bmatrix} u(0) - u^* \\ \gamma(0) - \gamma^* \end{bmatrix} + O \left( \frac{1}{\omega} + \max_i a_i \right),
\end{align*} \]
where \( u(t) = [u_1(t), \ldots, u_N(t)]^T \), \( u^* = [u_1^*, \ldots, u_N^*]^T \), \( \gamma(t) = [\gamma_{n+1}(t), \ldots, \gamma_N(t)]^T \), and \( \gamma^* = [1/D_{n+1,n+1}, \ldots, 1/D_{NN}]^T \).
Proof: Following the steps in the proofs of Theorem 2.1 and Theorem 2.3, one can obtain the following average error system, linearized about the equilibrium $(\tilde{u}^e, \tilde{\gamma}^e) = (0, 0)$:

$$
\frac{d}{d\tau} \begin{bmatrix} \delta \tilde{u}^{ave} \\ \delta \tilde{\gamma}^{ave} \end{bmatrix} = \frac{1}{\omega} \begin{bmatrix} \bar{A}_1 & 0 \\ 0 & \bar{A}_2 \end{bmatrix} \begin{bmatrix} \delta \tilde{u}^{ave} \\ \delta \tilde{\gamma}^{ave} \end{bmatrix},
$$

(2.67)

where $\delta \tilde{u}^{ave} = [(\delta \tilde{u})^1, \ldots, (\delta \tilde{u})^N]^T$, $\delta \tilde{\gamma}^{ave} = [(\delta \tilde{\gamma})^1, \ldots, (\delta \tilde{\gamma})^N]^T$.

$$
\bar{A}_1 = \begin{bmatrix} \bar{A}_{11} \\ \bar{A}_{12} \end{bmatrix},
$$

(2.68)

$$
\bar{A}_{11} = \begin{bmatrix} k_1 D_{11}^1 & k_1 D_{12}^1 & \cdots & k_1 D_{1n}^1 & k_1 D_{1N}^1 \\ k_2 D_{21}^1 & k_2 D_{22}^1 & \cdots & k_2 D_{2n}^1 & k_2 D_{2N}^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ k_n D_{n1}^1 & k_n D_{n2}^1 & \cdots & k_n D_{nn}^1 & k_n D_{nN}^1 \\ -k_{n+1} \frac{D_{n1,n+1}^{n+1}}{D_{n1,n+1}^{n+1}} & -k_{n+1} \frac{D_{n2,n+1}^{n+1}}{D_{n2,n+1}^{n+1}} & \cdots & -k_{n+1} \frac{D_{nN,n+1}^{n+1}}{D_{nN,n+1}^{n+1}} & -k_{n+1} \frac{D_{n1,n+1}^{n+1}}{D_{n1,n+1}^{n+1}} \end{bmatrix},
$$

(2.69)

$$
\bar{A}_{12} = \begin{bmatrix} -k_{n+1} \frac{D_{n1}^N}{D_{n1}^N} & -k_{n+1} \frac{D_{n2}^N}{D_{n2}^N} & \cdots & -k_{n+1} \frac{D_{nN}^N}{D_{nN}^N} & -k_{n+1} \frac{D_{nN}^N}{D_{nN}^N} \end{bmatrix},
$$

(2.70)

and $\bar{A}_2 = \text{diag}[-\sigma_n, \ldots, -\sigma_N]$. The $\bar{A}_{11}$ block of (2.68) is generated by the gradient-based players and the $\bar{A}_{12}$ block, by the Newton-based players.

As in Theorem 2.3, we require $\bar{A}_1$ and $\bar{A}_2$ to be Hurwitz for the origin of the average error system to be locally stable. The matrix $\bar{A}_2$ is Hurwitz since it is diagonal with negative elements. To analyze $A_1$, let

$$
\bar{K} = \text{diag} \left[ k_1, \ldots, k_n, -\frac{k_{n+1}}{D_{n1,n+1}^n}, \ldots, -\frac{k_N}{D_{nN}^n} \right]
$$

and recall that $D_{ii}^n < 0$. Then, we can write $\bar{A}_1$ as the matrix product

$$
\bar{A}_1 = \bar{K} D, \quad (2.71)
$$
where $D$ is the diagonally dominant matrix given by (2.16). From the Gershgorin Circle Theorem [48, Theorem 6.1.1], $\bar{A}_i$ is Hurwitz with eigenvalues contained in the set $\bigcup_{i=1}^{N} \rho_i$, where

$$\rho_i = \beta_i \left\{ z \in \mathbb{C} \left| \left| z - D_{ii}^i \right| < \sum_{j \neq i} |D_{ij}^i| \right. \right\},$$  

(2.72)

$$\beta_i = \begin{cases} k_i & \text{if } i \in \{1, \ldots, n\}, \\ \frac{k_i}{|D_{ii}^i|} & \text{if } i \in \{n + 1, \ldots, N\}. \end{cases}$$

(2.73)

The remainder of the proof follows from the proof of Theorem 2.1.

2.5 Summary

We have introduced a non-model based approach for convergence to the Nash equilibrium of static, noncooperative games with $N$ players and quadratic payoff functions. To determine their respective actions, the players need to measure only their own payoff values, which classifies this learning strategy as radically uncoupled, and for quadratic payoffs, convergence to $u^*$ is semi-global. Additionally, we have shown that this strategy approximates a player’s scaled best response dynamics and introduced a Newton-based Nash equilibrium seeking strategy so that this scaling is a known quantity.

If we assume the players possess a good estimate of $u^*$, based on either partial information of the game or historical data, this learning strategy remains attractive since it allows players to improve their initial actions by measuring only their payoff values and does not require the estimation of potentially highly uncertain parameters, e.g., competitors’ prices or consumer demand. Also, due to the dynamic nature of this algorithm, the players can track movements in $u^*$ should the game’s model change smoothly over time.


The dissertation author is the primary investigator and author of this work.
Chapter 3

Nash Equilibrium Seeking with Non-Quadratic Payoffs

To extend the results of Chapter 2, we consider $N$-player games with non-quadratic payoff functions that are the output of a dynamic system. Specifically, we consider general nonlinear differential equations with $N$ inputs and $N$ outputs, where in the steady state, the output signals represent the payoff functions of a noncooperative game played by the steady-state values of the input signals. When the players employ non-quadratic payoff functions, the game may possess multiple, isolated Nash equilibria. Hence, we pursue local convergence results because, in non-quadratic problems, global results come under strong restrictions. We show that the convergence is biased in proportion to the amplitudes of the perturbation signals and the third derivatives of the payoff functions, and confirm this convergence bias in a numerical example. In this example, we impose an action set $U \subset \mathbb{R}^N$ for the players and implement a Nash seeking strategy with projection to ensure that the players’ actions remain in $U$. This example is the first such application of extremum seeking with projection.

This chapter is organized as follows: In Section 3.1, we provide our most general result for $N$-player games with non-quadratic payoff functions and a dynamic mapping from the players’ actions to their payoff values. A detailed numerical example is provided in Section 3.2 before concluding with Section 3.3.
3.1 $N$-Player Games with Non-Quadratic Payoff Functions

Consider a more general noncooperative game with $N$ players and a dynamic mapping from the players’ actions $u_i$ to their payoff values $J_i$. Each player attempts to maximize the steady-state value of its payoff. Specifically, we consider a general nonlinear model,

$$\dot{x} = f(x, u),$$

$$J_i = h_i(x), \quad i \in \{1, \ldots, N\},$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^N$ is a vector of the players’ actions, $u_i$ is the action of player $i$, $J_i \in \mathbb{R}$ its payoff value, $f : \mathbb{R}^n \times \mathbb{R}^N \to \mathbb{R}^n$ and $h_i : \mathbb{R}^n \to \mathbb{R}$ are smooth, and $h_i$ is a possibly non-quadratic function. Oligopoly games may possess nonlinear demand and cost functions [75], which motivate the inclusion of the dynamic system (3.1) in the game structure and the consideration of non-quadratic payoff functions. For this scenario, we pursue local convergence results since the payoff functions $h_i$ may be non-quadratic and multiple, isolated Nash equilibria may exist. If the payoff functions are quadratic, semi-global practical stability can be achieved following the results of [103].

We make the following assumptions about this $N$-player game:

**Assumption 3.1** There exists a smooth function $l : \mathbb{R}^N \to \mathbb{R}^n$ such that

$$f(x, u) = 0 \quad \text{if and only if} \quad x = l(u).$$

**Assumption 3.2** For each $u \in \mathbb{R}^N$, the equilibrium $x = l(u)$ of (3.1) is locally exponentially stable.

Hence, we assume that for all actions, the nonlinear dynamic system is locally exponentially stable. We can relax the requirement that this assumption holds for each $u \in \mathbb{R}^N$ as we need to be only concerned with the action sets of the players, namely, $u \in U = U_1 \times \cdots \times U_N \subset \mathbb{R}^N$, and we do this in Section 3.2 for our numerical example. For notational convenience, we use this more restrictive case.
The following assumptions are central to our Nash seeking scheme as they ensure that at least one stable Nash equilibrium exists at steady state.

**Assumption 3.3** There exists at least one, possibly multiple, isolated stable Nash equilibria \( u^* = [u_1^*, \ldots, u_N^*] \) such that, for all \( i \in \{1, \ldots, N\} \),

\[
\frac{\partial (h_i \circ l)}{\partial u_i}(u^*) = 0, \quad \frac{\partial^2 (h_i \circ l)}{\partial u_i^2}(u^*) < 0.
\] (3.4)

**Assumption 3.4** The matrix,

\[
\Lambda = \begin{bmatrix}
\frac{\partial^2 (h_1 \circ l)(u^*)}{\partial u_1^2} & \frac{\partial^2 (h_1 \circ l)(u^*)}{\partial u_1 \partial u_2} & \cdots & \frac{\partial^2 (h_1 \circ l)(u^*)}{\partial u_1 \partial u_N} \\
\frac{\partial^2 (h_2 \circ l)(u^*)}{\partial u_1 \partial u_2} & \frac{\partial^2 (h_2 \circ l)(u^*)}{\partial u_2^2} & \cdots & \frac{\partial^2 (h_2 \circ l)(u^*)}{\partial u_2 \partial u_N} \\
\vdots & \vdots & & \vdots \\
\frac{\partial^2 (h_N \circ l)(u^*)}{\partial u_1 \partial u_N} & \frac{\partial^2 (h_N \circ l)(u^*)}{\partial u_N^2}
\end{bmatrix},
\] (3.5)

is strictly diagonally dominant and hence, nonsingular.

By Assumptions 3.3 and 3.4, \( \Lambda \) is Hurwitz.

As with the quadratic games considered in Chapter 2, each player converges to a neighborhood of \( u^* \) by implementing the extremum seeking strategy (2.9)–(2.10) to evolve its action \( u_i \) according to the measured value of its payoff \( J_i \). Unlike the previous games, however, we select the parameters \( k_i = \varepsilon \omega K_i = \mathcal{O}(\varepsilon \omega) \), where \( \varepsilon, \omega \) are small, positive constants and \( \omega \) is related to the players’ frequencies by (2.11). Intuitively, \( \omega \) is small since the players’ actions should evolve more slowly than the dynamic system, creating an overall system with two time scales. In contrast, our earlier analysis assumed \( 1/\omega \) to be small, which can be seen as the limiting case where the dynamic system is infinitely fast and allows \( \omega \) to be large.

Formulating the error system clarifies why these parameter selections are made. The error relative to the Nash equilibrium is denoted by (2.19), which in the time scale \( \tau = \omega t \), leads to

\[
\omega \frac{dx}{d\tau} = f(x, u^* + \tilde{u} + \mu(\tau)),
\] (3.6)

\[
\frac{d\tilde{u}_i}{d\tau} = \varepsilon K_i \mu_i(\tau) h_i(x), \quad i \in \{1, \ldots, N\},
\] (3.7)
where \( \tilde{u} = [\tilde{u}_1, \ldots, \tilde{u}_N] \), \( \mu(\tau) = [\mu_1(\tau), \ldots, \mu_N(\tau)] \), and \( \mu_i(\tau) = a_i \sin(\bar{\omega}_i \tau + \varphi_i) \).

The system (3.6)–(3.7) is in the standard singular perturbation form with \( \omega \) as a small parameter. Since \( \varepsilon \) is also small, we analyze (3.6)–(3.7) using the averaging theory for the quasi-steady state of (3.6), followed by the use of the singular perturbation theory for the full system.

### 3.1.1 Averaging Analysis

For the averaging analysis, we first “freeze” \( x \) in (3.6) at its quasi-steady state \( x = l(u^* + \tilde{u} + \mu(\tau)) \), which we substitute into (3.7) to obtain the “reduced system,”

\[
\frac{d\tilde{u}_i}{d\tau} = \varepsilon K_i \mu_i(\tau) (h_i \circ l)(u^* + \tilde{u} + \mu(\tau)).
\] (3.8)

This system is in the form to apply averaging theory [55] and leads to the following result:

**Theorem 3.1** Consider the system (3.8) for an \( N \)-player game under Assumptions 3.3 and 3.4, where \( \bar{\omega}_i \neq \bar{\omega}_j, \bar{\omega}_i \neq \bar{\omega}_j + \bar{\omega}_k, 2\bar{\omega}_i \neq \bar{\omega}_j + \bar{\omega}_k, \) and \( \bar{\omega}_i \neq 2\bar{\omega}_j + \bar{\omega}_k \) for all distinct \( i, j, k \in \{1, \ldots, N\} \) and \( \bar{\omega}_i \) is rational for all \( i \in \{1, \ldots, N\} \). There exist parameters \( \Xi, \xi > 0 \) and \( \varepsilon^* \), \( a^* \) such that, for all \( \varepsilon \in (0, \varepsilon^*) \) and \( a_i \in (0, a^*) \), if \( |\Theta(0)| \) is sufficiently small, then for all \( \tau \geq 0 \),

\[
|\Theta(\tau)| \leq \Xi e^{-\xi \tau} |\Theta(0)| + O \left( \varepsilon + \max_i a_i^3 \right),
\] (3.9)

where \( \Theta(\tau) = [\tilde{u}_1(\tau) - \sum_{j=1}^{N} c_{1j}^1 a_j^2, \ldots, \tilde{u}_N(\tau) - \sum_{j=1}^{N} c_{Nj}^N a_j^2]^T \), and

\[
\begin{bmatrix}
  c_{1j}^1 \\
  \vdots \\
  c_{Nj}^N
\end{bmatrix}
= -\frac{1}{4} \Lambda^{-1} \mathbf{g}_j,
\] (3.10)
\[ \mathbf{g}_j \triangleq \begin{bmatrix} \frac{\partial^3 (h_1 \circ l)(u^*)}{\partial u_1 \partial u_j^3} \\
 \vdots \\
 \frac{\partial^3 (h_{j-1} \circ l)(u^*)}{\partial u_{j-1} \partial u_j^3} \\
 \frac{\partial^3 (h_j \circ l)(u^*)}{\partial u_j^3} \\
 \frac{\partial^3 (h_{j+1} \circ l)(u^*)}{\partial u_j \partial u_{j+1}} \\
 \vdots \\
 \frac{\partial^3 (h_N \circ l)(u^*)}{\partial u_j \partial u_N} \end{bmatrix}, \tag{3.11} \]

Proof: As already noted, the form of (3.8) allows for the application of averaging theory, which yields the average error system,

\[
\frac{d\tilde{u}_i^{\text{ave}}}{d\tau} = \varepsilon \left( \frac{K_i}{T} \int_0^T \mu_i(\tau) (h_i \circ l)(u^* + \tilde{u}^{\text{ave}} + \mu(\tau)) \, d\tau \right). \tag{3.12}
\]

The equilibrium \( \tilde{u}^e = [\tilde{u}_1^e, \ldots, \tilde{u}_N^e] \) of (3.12) satisfies

\[
0 = \frac{1}{T} \int_0^T \mu_i(\tau) (h_i \circ l)(u^* + \tilde{u}^e + \mu(\tau)) \, d\tau, \tag{3.13}
\]

for all \( i \in \{1, \ldots, N\} \), and we postulate that \( \tilde{u}^e \) has the form,

\[
\tilde{u}_i^e = \sum_{j=1}^N b_j^i a_j + \sum_{j=1}^N \sum_{k \geq j} c_{jk}^i a_j a_k + O \left( \max_i a_i^3 \right). \tag{3.14}
\]

By approximating \( (h_i \circ l) \) about \( u^* \) in (3.13) with a Taylor polynomial and substituting (3.14), the unknown coefficients \( b_j^i \) and \( c_{jk}^i \) can be determined.

To capture the effect of higher order derivatives on average error system’s equilibrium, we use the Taylor polynomial approximation [2], which requires \( (h_i \circ l) \) to be \( k + 1 \) times differentiable,

\[
(h_i \circ l)(u^* + \tilde{u}^e + \mu(\tau))
= \sum_{|\alpha| = 0}^k \frac{D^\alpha (h_i \circ l)(u^*)}{\alpha!} (\tilde{u}^e + \mu(\tau))^\alpha + \sum_{|\alpha| = k+1}^k \frac{D^\alpha (h_i \circ l)(\zeta)}{\alpha!} (\tilde{u}^e + \mu(\tau))^\alpha,
\]

\[
= \sum_{|\alpha| = 0}^k \frac{D^\alpha (h_i \circ l)(u^*)}{\alpha!} (\tilde{u}^e + \mu(\tau))^\alpha + O \left( \max_i a_i^{k+1} \right), \tag{3.15}
\]

where \( \zeta \) is a point on the line segment that connects the points \( u^* \) and \( u^* + \tilde{u}^e + \mu(\tau) \). In (3.15), we have used multi-index notation, namely, \( \alpha = (\alpha_1, \ldots, \alpha_N) \),
\[ |\alpha| = \alpha_1 + \cdots + \alpha_N, \; \alpha! = \alpha_1! \cdots \alpha_N!, \; u^\alpha = u_1^{\alpha_1} \cdots u_N^{\alpha_N}, \; \text{and} \; D^\alpha(h_i \circ l) = \partial^{\alpha_i}(h_i \circ l)/\partial u_1^{\alpha_1} \cdots \partial u_N^{\alpha_N}. \] 

The second term on the last line of (3.15) follows by substituting the postulated form of \( \tilde{u}^e \) (3.14).

For this analysis, we choose \( k = 3 \) to capture the effect of the third order derivative on the system as a representative case. Higher-order estimates of the bias can be pursued if the third order derivative is zero. Substituting (3.15) into (3.13) and computing the average of each term gives

\[
0 = \frac{a_2^2}{2} \sum_{j=1}^N \partial^2(h_i \circ l)(u^*) + \sum_{j \neq i} \partial^2(h_i \circ l)(u^*) + 4 \partial^2(h_i \circ l)(u^*)
\]

where we have noted (3.4), utilized (3.14), and computed the integrals shown in Sections A.1 and A.2 of Appendix A. Substituting (3.14) into (3.16) and matching first order powers of \( a_i \) gives

\[
\begin{bmatrix}
0 \\
 \vdots \\
0
\end{bmatrix}
= a_1 \Lambda 
\begin{bmatrix}
b_1^1 \\
 \vdots \\
b_1^N
\end{bmatrix}
+ \cdots + a_N \Lambda 
\begin{bmatrix}
b_N^1 \\
 \vdots \\
b_N^N
\end{bmatrix}
\]

which implies that \( b_{ij}^i = 0 \) for all \( i, j \) since \( \Lambda \) is nonsingular by Assumption 3.4. Similarly, matching second order terms of \( a_i \), and substituting \( b_{ij}^i = 0 \) to simplify the resulting expressions, yields

\[
\begin{bmatrix}
0 \\
 \vdots \\
0
\end{bmatrix}
= \sum_{j=1}^N \sum_{k>j} a_j a_k \Lambda 
\begin{bmatrix}
c_1^1 \\
 \vdots \\
c_1^N
\end{bmatrix}
+ \sum_{j=1}^N a_j^2 \Lambda 
\begin{bmatrix}
c_1^1 \\
 \vdots \\
c_1^N
\end{bmatrix}
+ \frac{1}{4} g_j,
\]

where \( g_j \) is defined in (3.11). Thus, \( c_{jk}^i = 0 \) for all \( i, j, k \) when \( j \neq k \), and \( c_{jj}^i \) is given by (3.10) for \( i, j \in \{1, \ldots, N\} \). The equilibrium of the average system is
then
\[ \hat{u}_i^e = \sum_{j=1}^{N} c_{jj} a_j^2 + O \left( \max_i a_i^3 \right). \] (3.17)

By again utilizing a Taylor polynomial approximation, one can show that the Jacobian \( \Psi^{\text{ave}} = [\psi_{i,j}]_{N \times N} \) of (3.12) at \( \tilde{u}^e \) has elements given by
\[ \psi_{i,j} = \frac{\varepsilon K_i}{T} \int_0^T \mu_i(\tau) \frac{\partial (h_i \circ l)}{\partial u_j} (u^* + \tilde{u}^e + \mu(\tau)) \, d\tau, \]
\[ = \frac{1}{2} \varepsilon K_i a_i^2 \frac{\partial^2 (h_i \circ l)}{\partial u_i \partial u_j} (u^*) + O \left( \varepsilon \max_i a_i^3 \right). \] (3.18)

By Assumptions 3.3 and 3.4, \( \Psi^{\text{ave}} \) is Hurwitz for sufficiently small \( a_i \), which implies that the equilibrium (3.17) of the average error system (3.12) is locally exponentially stable, i.e., there exist constants \( \Xi, \xi > 0 \) such that
\[ |\tilde{u}^{\text{ave}}(\tau) - \tilde{u}^e_\tau | \leq \Xi e^{-\xi \tau} |\tilde{u}^{\text{ave}}(0) - \tilde{u}^e|, \]
where with [55, Theorem 10.4] implies
\[ |\tilde{u}(\tau) - \tilde{u}^e| \leq \Xi e^{-\xi \tau} |\tilde{u}(0) - \tilde{u}^e| + O \left( \varepsilon \right), \] (3.19)
provided \( \tilde{u}(0) \) is sufficiently close to \( \tilde{u}^{\text{ave}}(0) \). Defining \( \Theta(\tau) \) as in Theorem 3.1 completes the proof.

From Theorem 3.1, we see that \( u \) of reduced system (3.8) converges to a region that is biased away from the Nash equilibrium \( u^* \). This bias is in proportion to the perturbation magnitudes \( a_i \) and the third derivatives of the payoff functions, which are captured by the coefficients \( c_{jj} \). Specifically, \( \hat{u}_i \) of the reduced system converges to \( u_i^* + \sum_{j=1}^{N} c_{jj} a_j^2 + O(\varepsilon + \max_i a_i^3) \) as \( t \to \infty \).

Theorem 3.1 can be viewed as a generalization of Theorem 2.1, but with a focus on the error system to highlight the effect of the payoff functions’ non-quadratic terms on the players’ convergence. This emphasis is needed because \( u_i \) of the reduced system converges to an \( O(\varepsilon + \max_i a_i) \)-neighborhood of \( u_i^* \), as in the quadratic payoff case, since \( u_i = \hat{u}_i + \mu_i \).

### 3.1.2 Singular Perturbation Analysis

We analyze the full system (3.6)–(3.7) in the time scale \( \tau = \omega t \) using singular perturbation theory [55]. In Section 3.1.1, we analyzed the reduced model (3.8)
and now, must study the boundary layer model to state our convergence result. First, however, we translate the equilibrium of the reduced model to the origin by defining \( z_i = \tilde{u}_i - \tilde{u}_i^p \), where by [55, Theorem 10.4], \( \tilde{u}_i^p \) is a unique, exponentially stable, \( T \)-periodic solution \( \tilde{u}^p = [\tilde{u}_1^p, \ldots, \tilde{u}_N^p] \) such that

\[
\frac{d\tilde{u}^p_i}{d\tau} = \varepsilon K_i \mu_i(\tau)(u^* + \tilde{u}^p + \mu(\tau)).
\]

(3.20)

In this new coordinate system, we have for \( i \in \{1, \ldots, N\} \)

\[
\frac{dz_i}{d\tau} = \varepsilon K_i \mu_i(\tau) \left[ (h_i \circ l)(u^* + \tilde{u}^p + \mu(\tau)) \right]
\]

(3.21)

\[
\omega \frac{dx}{d\tau} = f(x, u^* + z + \tilde{u}^p + \mu(\tau)),
\]

(3.22)

which from Assumption 3.1 has the quasi-steady state \( x = l(u^* + z + \tilde{u}^p + \mu(\tau)) \), and consequently, the reduced model in the new coordinates is

\[
\frac{dz_i}{d\tau} = \varepsilon K_i \mu_i(\tau) \left[ (h_i \circ l)(u^* + z + \tilde{u}^p + \mu(\tau)) \right]
\]

\[
- (h_i \circ l)(u^* + \tilde{u}^p + \mu(\tau)),
\]

(3.23)

which has an equilibrium at \( z = 0 \) that is exponentially stable for sufficiently small \( a_i \).

To formulate the boundary layer model, let \( y = x - l(u^* + z + \tilde{u}^p + \mu(\tau)) \), and then in the time scale \( t = \tau/\omega \), we have

\[
\frac{dy}{dt} = f(y + l(u^* + z + \tilde{u}^p + \mu(\tau)), u^* + z + \tilde{u}^p + \mu(\tau)),
\]

\[
= f(y + l(u), u),
\]

(3.24)

where \( u = u^* + \tilde{u} + \mu(\tau) \) should be viewed as a parameter independent of the time variable \( t \). Since \( f(l(u), u) = 0 \), \( y = 0 \) is an equilibrium of (3.24) and is exponentially stable by Assumption 3.2.

With \( \omega \) as a singular perturbation parameter, we apply Tikhonov’s Theorem on the Infinite Interval [55, Theorem 11.2] to (3.21)–(3.22), which requires the origin to be an exponentially stable equilibrium point of both the reduced model (3.23) and the boundary layer model (3.24) and leads to the following:

- the solution \( z(\tau) \) of (3.21) is \( O(\omega) \)-close to the solution \( \bar{z}(\tau) \) of the reduced model (3.23), so
the solution \( \tilde{u}(\tau) \) of (3.7) converges exponentially to an \( O(\omega) \)-neighborhood of the \( T \)-periodic solution \( \tilde{u}^p(\tau) \), and

- the \( T \)-periodic solution \( \tilde{u}^p(\tau) \) is \( O(\varepsilon) \)-close to the equilibrium \( \tilde{u}^e \).

Hence, as \( t \to \infty \), \( \tilde{u}(\tau) \) converges to an \( O(\varepsilon) \)-neighborhood of \( \tilde{u}^e \) = \( \left[ \sum_{j=1}^N c_{i,j}a_j^2, \ldots, \sum_{j=1}^N c_{N,j}a_j^2 \right] + O(\max_i a_i^2) \). Since \( u(\tau) - u^* = \tilde{u}(\tau) + \mu(\tau) = \tilde{u}(\tau) + O(\max_i a_i), u(\tau) \) converges to an \( O(\omega + \varepsilon + \max_i a_i) \)-neighborhood of \( u^* \).

Also from Tikhonov’s Theorem on the Infinite Interval, the solution \( x(\tau) \) of (3.22), which is the same as the solution of (3.6), satisfies

\[
x(\tau) - l(u^* + \tilde{u}^r(\tau) + \mu(\tau)) - y(t) = O(\omega),
\]

where \( \tilde{u}^r(\tau) \) is the solution of the reduced model (3.8) and \( y(t) \) is the solution of the boundary layer model (3.24). Rearranging terms and subtracting \( l(u^*) \) from both sides yields

\[
x(\tau) - l(u^*) = O(\omega) + l(u^* + \tilde{u}^r(\tau) + \mu(\tau)) - l(u^*) + y(t).
\]

After noting that

- \( \tilde{u}^r(\tau) \) converges exponentially to \( \tilde{u}^p(\tau) \), which is \( O(\varepsilon) \)-close to the equilibrium \( \tilde{u}^e \),

- \( \mu(\tau) \) is \( O(\max_i a_i) \), and

- \( y(t) \) is exponentially decaying,

we conclude that \( x(\tau) - l(u^*) \) exponentially converges to an \( O(\omega + \varepsilon + \max_i a_i) \)-neighborhood of the origin. Thus, \( J_i = h_i(x) \) exponentially converges to an \( O(\omega + \varepsilon + \max_i a_i) \)-neighborhood of the payoff value \((h_i \circ l)(u^*)\).

We summarize with the following theorem:

**Theorem 3.2** Consider the system (3.1)–(3.2) with (2.9)–(2.10) for an \( N \)-player game under Assumptions 3.1–3.4, where \( \omega_i \neq \omega_j, \omega_i \neq \omega_j + \omega_k, 2\omega_i \neq \omega_j + \omega_k, \) and \( \omega_i \neq 2\omega_j + \omega_k \) for all distinct \( i, j, k \in \{1, \ldots, N\} \) and where \( \omega_i/\omega_j \) is rational for all \( i, j \in \{1, \ldots, N\} \). There exists \( \omega^* > 0 \) and for any \( \omega \in (0, \omega^*) \) there exist
\( \varepsilon^*, a^* > 0 \) such that for the given \( \omega \) and any \( \varepsilon \in (0, \varepsilon^*) \) and \( \max_i a_i \in (0, a^*) \), the solution \((x(t), u_1(t), \ldots, u_N(t))\) converges exponentially to an \( O(\omega + \varepsilon + \max_i a_i) \)-neighborhood of the point \((l(u^*), u_1^*, \ldots, u_N^*)\), provided the initial conditions are sufficiently close to this point.

### 3.2 Example

For an example non-quadratic game with players that employ the extremum seeking strategy (2.9)–(2.10), we consider the system,

\[
\begin{align*}
\dot{x}_1 &= -4x_1 + x_1 x_2 + u_1, \\
\dot{x}_2 &= -4x_2 + u_2, \\
J_1 &= -16x_1^2 + 8x_1x_2 - x_1^2 x_2^2 - 6x_1 x_2^2 + \left(24 + \frac{5}{32}\right)x_1 x_2 - \frac{5}{8}x_1, \\
J_2 &= -64x_2^3 + 48x_1 x_2^2 - 12x_1 x_2^2,
\end{align*}
\]

whose equilibrium state is given by \((\bar{x}_1, \bar{x}_2) = (4u_1/(16 - u_2), u_2/4)\). The Jacobian at the equilibrium \((\bar{x}_1, \bar{x}_2)\) is Hurwitz if \(u_2 < 16\). Thus, the \((\bar{x}_1, \bar{x}_2)\) is locally exponentially stable, but not for all \((u_1, u_2) \in \mathbb{R}^2\), violating Assumption 3.2. However, as noted earlier, this restrictive requirement of local exponential stability for all \(u \in \mathbb{R}^N\) was done merely for notational convenience, and we actually only require this assumption to hold for the players’ action sets. Hence, in this example, we restrict the players’ actions to the set \(U = \{(u_1, u_2) \in \mathbb{R}^2 \mid u_1, u_2 \geq 0, u_2 < 16\}\).

At \(x = \bar{x}\), the payoff functions are

\[
J_1 = -u_1^2 + \frac{3}{2} u_1 u_2 - \frac{5}{32} u_1, \quad J_2 = -u_2^3 + 3u_1 u_2, \quad (3.27)
\]

which have the associated reaction curves, defined on the action set \(U\), shown below:

\[
l_1(u_2) = \begin{cases} 
\frac{3}{4}u_2 - \frac{5}{64} & \text{if } \frac{5}{48} \leq u_2 < 16, \\
0 & \text{if } 0 \leq u_2 < \frac{5}{48},
\end{cases} \quad (3.28)
\]

\[
l_2(u_1) = \sqrt{u_1}. \quad (3.29)
\]
Figure 3.1: The steady-state payoff function surfaces (3.27) with both their associated reaction curves (black) (3.28)–(3.29), which lie in the action set $U$, and the extremals $\partial J_i/\partial u_i = 0$ (dashed yellow), which lie outside of $U$, superimposed.

Figure 3.1 depicts each player’s payoff surface with both the reaction curves and the extremals $\partial J_i/\partial u_i = 0$ that lie outside of $U$ superimposed. These reaction curves have two interch3-sections in the interior of $U$ (Figure 3.2(a)), which correspond to two Nash equilibria: $(u^*_1, u^*_2) = (25/64, 5/8)$ and $(v^*_1, v^*_2) = (1/64, 1/8)$. At $u^*$, Assumptions 3.3 and 3.4 are satisfied, implying the stability of $u^*$, whereas at $v^*$, these assumptions are violated—as they must be—since further analysis will show that $v^*$ is an unstable Nash equilibrium. The reaction curves also interch3-sect on the action set boundary $\partial U$ at $(0, 0)$, which means this game has a Nash equilibrium on the boundary that the players may seek depending on the game’s initial conditions. To ensure the players remain in the action set, we employ a modified Nash seeking strategy that utilizes projection [59]. For this example, we will first present simulation results for Nash seeking in the interior of $U$ before considering the Nash equilibrium on the boundary.

3.2.1 Nash Seeking in the Interior

For the Nash equilibrium points in the interior of $U$, we want to determine their stability and to quantify the convergence bias due to the non-quadratic payoff function. First, we compute the average error system for the reduced model
Figure 3.2: (a) The reaction curves associated with the steady-state payoff functions (3.27) with the stable Nash equilibrium (green circle), the unstable Nash equilibrium (red square), and the Nash equilibrium on the boundary (blue star), and (b) the convergence bias (+) relative to each interior Nash equilibrium, due to the non-quadratic payoff function of player 2 and lies along the reaction curve of player 1.

According to (3.12), obtaining

\[
\frac{d\bar{u}^{\text{ave}}_1}{d\tau} = \varepsilon K_1 a_1^2 \left( -\bar{u}^{\text{ave}}_1 + \frac{3}{4} \bar{u}^{\text{ave}}_2 \right),
\]
\[
\frac{d\bar{u}^{\text{ave}}_2}{d\tau} = \varepsilon K_2 a_2^2 \left( \frac{3}{2} \bar{u}^{\text{ave}}_1 - 3\eta^* \bar{u}^{\text{ave}}_2 - \frac{3}{2} (\bar{u}^{\text{ave}}_2)^2 - \frac{3}{8} a_2^2 \right),
\]

where \( \eta^* = u^*_2 \) or \( v^*_2 \). The equilibria of (3.30)–(3.31) are given by \((\bar{u}^{e}_1(\eta^*), \bar{u}^{e}_2(\eta^*)) = \left((9 - 24\eta^*)/32 \pm 3\sqrt{(3 - 8\eta^*)^2 - 16a_2^3/32}, 4\bar{u}^{e}_1(\eta^*)/3 \right)\). While the system appears to have four equilibria, two for each value of \( \eta^* \), two equilibria correspond to the difference between the two Nash equilibrium points, meaning that in actuality, only two equilibria exist and can be written as

\[
(\bar{u}^{e}_1(\eta^*), \bar{u}^{e}_2(\eta^*)) = \left( \delta(\eta^*), \frac{4}{3} \delta(\eta^*) \right),
\]

where \( \delta(\eta^*) = (9 - 24\eta^*)/32 - \text{sgn}(3 - 8\eta^*)3\sqrt{(3 - 8\eta^*)^2 - 16a_2^3/32} \) is the smallest value of \( \bar{u}^{e}_1(\eta^*) \) for each Nash equilibrium. In the sequel, we omit the dependency of \( \bar{u}^{e}_1, \bar{u}^{e}_2, \) and \( \delta \) on \( \eta^* \) for conciseness. As seen in Figure 3.2(b), the equilibrium is biased away from the reaction curve of player 2 but still lies on the reaction curve \( l_1 \) of player 1 since only the payoff function for player 2 is non-quadratic.
The Jacobian $\Psi^{\text{ave}}$ of the average error system for the reduced model is

$$
\Psi^{\text{ave}} = \begin{bmatrix}
-\kappa_1 & \frac{4}{3}\kappa_1 \\
\frac{2}{3}\kappa_2 & -\kappa_2 \left(4\delta + 3\eta^*\right)
\end{bmatrix}
$$

when evaluated at $(\tilde{u}_1^e, \tilde{u}_2^e) = (\delta, 4\delta/3)$, where $\kappa_1 = \varepsilon K_1 a_1^2$ and $\kappa_2 = \varepsilon K_2 a_2^2$. Its characteristic equation is

$$
\lambda^2 + \left(\kappa_1 + \kappa_2 \left(4\delta + 3\eta^*\right)\right)\lambda + \kappa_1 k_2 \left(4\delta + 3\eta^* - 9/8\right) = 0.
$$

Thus, $\Psi^{\text{ave}}$ is Hurwitz if and only if $\zeta_1$ and $\zeta_2$ are positive. For sufficiently small $a_2$ so that $\delta \approx 0$, $\zeta_1, \zeta_2 > 0$ when $\eta^* = u_2^* = 5/8$, and $\zeta_2 < 0$ when $\eta^* = v_2^* = 1/8$, which implies $u^*$ is a stable Nash equilibrium and $v^*$ is unstable. Closer analysis shows that the Jacobian associated with $v^*$ has two real eigenvalues—one that is negative and one that is positive.

For the simulations, we select $k_1 = k_2 = 1.75$, $a_1 = 0.06$, $a_2 = 0.03$, $\omega_1 = 1.2$, and $\omega_2 = 2$, where the parameters are chosen to be small, in particular the perturbation frequencies $\omega_i$, since the perturbation must occur at a time scale that is slower than the fast time scale of the nonlinear system. Figures 3.3(a) and 3.3(b) depict the evolution of the players’ actions $\hat{u}_1$ and $\hat{u}_2$ initialized at $(u_1(0), u_2(0)) = (\hat{u}_1(0), \hat{u}_2(0)) = (0.2, 0.9)$ and $(0.1, 0.15)$. The state $(x_1, x_2)$ is
Figure 3.4: Phase portrait of the $\hat{u}$-reduced model average system, with the players’ $\hat{u}$-trajectories superimposed for each initial condition. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and Nash equilibrium on $\partial U$ (blue star) are denoted. The shaded region is the set difference $U \setminus \hat{U}$. If $\hat{u}_i \in \hat{U}$, then $u_i \in U$. The zoomed-in area in shows the players’ convergence to an almost-periodic orbit that is biased away from the stable Nash equilibrium and centered on the reaction curve $l_1$. The average actions of the players are marked by $\times$.

initialized at the origin in both cases. We show $\hat{u}_i$ instead of $u_i$ to better illustrate the convergence of the players’ actions to a neighborhood near—but biased away from—the Nash strategies since $u_i$ contains the additive signal $\mu_i(t)$. In both scenarios, the average actions of both player 1 and player 2 lie below $u^*$, which is consistent with the equilibrium of the reduced model’s average error system (3.32).

The slow initial convergence in Figure 3.3(b) can be explained by examining the phase portrait of the average of the reduced model $\hat{u}$-system. The initial condition $(0.1, 0.15)$ lies near the unstable equilibrium, causing the slow initial convergence seen in Figure 3.3(b). In Figure 3.4, the stable and unstable interior Nash equilibria (green circle and red square), the Nash equilibrium on $\partial U$ (blue star), and the eigenvectors (magenta arrows) associated with each interior Nash equilibrium are denoted. The eigenvectors are, in general, neither tangent nor perpendicular to the reaction curves at the Nash equilibrium points. The phase portrait and orientation of the eigenvectors do depend on the values of $a_i$ and $k_i$, but the stability of each Nash equilibrium is invariant to their values, provided these parameters are positive and small as shown in Sections 2.2 and 3.1. Since
Figure 3.4 is a phase portrait of the \( \hat{u} \) -reduced system, we include the shaded regions to show that \( \hat{u}_i \) must be restricted to the set \( \hat{U} = \{(\hat{u}_1, \hat{u}_2) \in \mathbb{R}^2 \mid a_1 \leq \hat{u}_1, a_2 \leq \hat{u}_2 \leq 16 - a_2\} \) to ensure that \( u_i \) lies in \( U \). Hence, the shaded region’s outer edge is \( \partial U \), the inner edge is \( \partial \hat{U} \), and the shaded interior is the set difference \( U \setminus \hat{U} \).

When the players’ initial conditions lie in the stable region of the phase portrait, the players’ actions remain in \( U \) and converge to a neighborhood of the stable Nash equilibrium in the interior of \( U \). The zoomed-in area highlights the convergence of the trajectories to an almost-periodic orbit that is biased away from the Nash equilibrium \( u^* \). The depicted orbit consists of the last 300 sec of each trajectory with the players’ average actions denoted by \( \times \). These average actions lie on the reaction curve \( l_1 \) as predicted by (3.32).

### 3.2.2 Nash Seeking with Projection

To ensure that a player’s action remains in the action set \( U \), we employ a modified extremum seeking strategy that utilizes projection. Namely, the players implement the following strategy:

\[
\begin{align*}
\dot{\hat{u}}_i(t) &= k_i \text{Proj}\{\mu_i(t)J_i(t); a_i\}, \\
u_i(t) &= \hat{u}_i(t) + \mu_i(t),
\end{align*}
\]

(3.33) (3.34)

where \( \mu_i(t) \) is defined as in (2.9)–(2.10) and

\[
\text{Proj}\{\phi; \chi\} = \begin{cases} 
0 & \text{if } \hat{u}_i \leq \chi \text{ and } \phi < 0, \\
\phi & \text{otherwise.}
\end{cases}
\]

(3.35)

**Lemma 3.1** The following properties hold for the projection operator (3.35):

1. For all \( (\phi, \hat{u}_i) \in \mathbb{R}^2 \), \( |\text{Proj}\{\phi; \chi\}| \leq |\phi| \).

2. With \( \hat{u}_i(0) \geq \chi \), the update law (3.33) with the projection operator (3.35) guarantees that \( \hat{u}_i \) is maintained in the projection set, namely, \( \hat{u}_i(t) \in [\chi, \infty) \) for all \( t \geq 0 \).
3. For all \((\phi, \hat{u}_i) \in \mathbb{R}^2\) and \(u_i^* \geq \chi\), the following holds:

\[
(u_i^* - \hat{u}_i)(\phi - \text{Proj}\{\phi; \chi\}) \leq 0.
\] (3.36)

Proof: Points 1 and 2 are immediate. Point 3 follows from the calculation,

\[
(u_i^* - \hat{u}_i)(\phi - \text{Proj}\{\phi; \chi\}) = \begin{cases} 
(u_i^* - \hat{u}_i)\phi & \text{if } \hat{u}_i \leq \chi \text{ and } \phi < 0, \\
0 & \text{otherwise}
\end{cases}
\] (3.37)

since \(u_i^* - \hat{u}_i \geq 0\) when \(u_i \leq \chi\).

Several challenges remain in the development of convergence proofs for Nash seeking players with projection, such as computing the average error system and defining the precise notion of stability for equilibria that exist on the action space boundaries. To alleviate concerns of existence and uniqueness of solutions that arise from the projection operator being discontinuous, a Lipschitz continuous version of projection [59] can be used.

Figure 3.5 depicts the trajectories \(\hat{u}_1\) and \(\hat{u}_2\) when the players employ (3.33)–(3.34) with the same parameters as in Section 3.2.1 and the initial condition \((u_1(0), u_2(0)) = (\hat{u}_1(0), \hat{u}_2(0)) = (0.2, 0.03)\). For this initial condition, the players reach \(\partial \hat{U}\) where \(\hat{u}_1 = a_1 = 0.06\) while \(\hat{u}_2\) increases until the trajectory reaches the eigenvector and \(\hat{u}_1\) is pushed away from the boundary. From this point, the players follow a trajectory that is similar to the one depicted in Figure 3.3(b).

While the modified strategy (3.33)–(3.34) ensures that the players’ actions lie in \(U\), it effectively prevents the players from converging to a neighborhood of any equilibria that lie in \(U \setminus \hat{U}\). To address this limitation, we propose the modified projection strategy,

\[
\dot{u}_i(t) = k_i \text{Proj}\{\mu_i(t)J_i(t); 0\},
\]
(3.38)

\[
u_i(t) = \dot{u}_i(t) + \min(1, |\dot{u}_i(t)|/a_i)\mu_i(t),
\]
(3.39)

where \(\text{Proj}\{\phi; \chi\}\) is defined as in (3.35). The strategy (3.38)–(3.39) is equivalent to (3.33)–(3.34) when \(\chi = a_i\) since \(\text{Proj}\{\mu_i(t)J_i(t); a_i\}\) ensures \(\min(1, |\dot{u}_i|/a_i) = 1\).

In Figure 3.6, the players employ (3.38)–(3.39) for the same scenario shown in Figure 3.5. With this strategy, the players approach the boundary \(u_1 = \hat{u}_1 = 0\),
Figure 3.5: (a) Time history of the two-player game initialized at (0.2, 0.03) when implementing the modified Nash seeking strategy (3.33)–(3.34), and (b) planar \( \hat{u} \)-trajectories superimposed on the \( \hat{u} \)-reduced model average system phase portrait. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and the Nash equilibrium on \( \partial U \) (blue star) are denoted. The shaded region is the set difference \( U \setminus \hat{U} \).

where, unlike at the boundary \( \hat{u}_1 = 0.06 \), the vector field causes \( \hat{u}_2 \) to decrease toward the Nash equilibrium at \( (0, 0) \). This Nash equilibrium is attainable only by allowing \( \hat{u}_i \) to enter \( U \setminus \hat{U} \), but the convergence rate slows near the boundary since \( \min(a_i, |\hat{u}_i|)/a_i \to 0 \) as \( \hat{u}_i \) approaches \( \partial U \) and due to the nearby unstable Nash equilibrium (Figure 3.6). If the nonnegative action requirement were relaxed to allow the perturbation to leave \( U \), i.e., we restrict \( u_i \) to the set \( \hat{U} = \{(u_1, u_2) \in \mathbb{R}^2 \mid -a_1 \leq u_1, -a_2 \leq u_2 < 16\} \), then the point \( (0, 0) \) would be attainable using (3.33)–(3.34) with \( \chi = 0 \). This strategy would exhibit a faster convergence rate along \( \partial U \) than if the players used (3.38)–(3.39), but convergence would still be slow due to the presence of the unstable Nash equilibrium.

### 3.3 Summary

We have extended the Nash equilibrium seeking strategy for static, noncooperative games with \( N \) players to games with non-quadratic functions and games where the players’ actions serve as inputs to a general, stable nonlinear differ-
Figure 3.6: (a) Time history of the two-player game initialized at (0.2, 0.03) when implementing the modified Nash seeking strategy (3.38)–(3.39), and (b) planar \( \hat{u} \)-trajectories superimposed on the \( \hat{u} \)-reduced model average system phase portrait. The unstable Nash equilibrium (red square), its eigenvectors (magenta arrows), and the Nash equilibrium on \( \partial U \) (blue star) are denoted. The shaded region is the set difference \( U \setminus \hat{U} \).


The dissertation author is the primary investigator and author of this work.
Chapter 4

Nash Equilibrium Seeking with Infinitely-Many Players

We study here the problem of computing, in real time, the Nash equilibria of static noncooperative games with infinitely-many players with quadratic payoff functions. In games of this type, the action of a single player cannot affect the outcome of the game, and yet, by utilizing extremum seeking, a non-model based approach, the players achieve stable, local attainment of their Nash strategies by measuring only their own payoff values. The players are not required to know the mathematical model of their payoff function or the game. The tradeoff is that convergence in this case is proved only locally (or at best semi-globally, see [103]). Economic models with a continuum of players have been studied since the 1960s [4, 43, 95].

This chapter is organized as follows: Section 4.1 introduces Nash equilibrium seeking for games with uncountably many players for two classes of quadratic payoff functions, Section 4.2 presents our convergence results, Section 4.3 provides a numerical example, and Section 4.4 summarizes the result.

4.1 Nash Equilibrium Seeking

We consider static games with uncountably many (non-atomic) players that wish to maximize their quadratic payoff functions. For such games, we associate
with each player a point \( x \) in the unit interval \([0, 1]\) and denote the action of player \( x \) by \( u(x) \) and its payoff value by \( J(x) \).

By utilizing extremum seeking, a player can stably attain a Nash equilibrium \( u^*(x) \) by evolving its action \( u(x) \) according to its measured payoff value \( J(x) \). Specifically, the players employ the time-varying strategy

\[
\frac{\partial}{\partial t} \dot{u}(x, t) = k(x) \mu(x, t) J(x, t),
\]

(4.1)

\[
u(x, t) = \dot{u}(x, t) + \mu(x, t),
\]

(4.2)

where \( \mu(x, t) = a(x) \sin(\omega(x) t + \varphi(x)) \), \( a(x) \) is measurable, positive, and bounded for all \( x \in [0, 1] \), and \( \omega(x), k(x) > 0 \) for all \( x \in [0, 1] \). The strategy (4.1)–(4.2) requires player \( x \) to know only its payoff value \( J(x) \). Knowledge of the mathematical form of the payoff functions or of the other players’ actions is not needed. Figure 4.1 depicts a noncooperative game played by infinitely many players implementing the Nash seeking strategy (4.1)–(4.2).

We study two classes of quadratic payoff functions. First, we analyze payoff functions of the form

\[
J(x, t) = -c(x) u^2(x, t) + d(x) u(x, t) \int_0^1 p(y) u(y, t) dy + q(x) u(x, t) + \int_0^1 r(x, y) u(y, t) dy + e(x),
\]

(4.3)

where \( c, d, p, q, r, e \) are all measurable and bounded functions. Moreover, \( c(x) > 0 \) for all \( x \in [0, 1] \), and \( \int_0^1 p(y) d(y)/c(y) dy \neq 2 \) is assumed because the Nash equilibrium for games with payoff functions of the form (4.3) is

\[
u^*(x) = \frac{d(x) \int_0^1 \frac{p(y) q(y)}{c(y)} dy}{c(x) \left( 4 - 2 \int_0^1 \frac{p(y) d(y)}{c(y)} dy \right)} + \frac{1}{2} \frac{q(x)}{c(x)}.
\]

(4.4)

The other class of quadratic payoff functions that we consider, which is not a subset of (4.3), is

\[
J(x, t) = -u^2(x, t) + 2u(x, t) \int_0^1 r(x, y) u(y, t) dy,
\]

(4.5)

where \( r \) is measurable, bounded, and cannot be expressed as a product of two single-argument functions, i.e., \( r(x, y) \neq g(x) h(y) \) in general. (Compare the second
term of (4.5) to the second term of (4.3).) The payoff functions (4.5) yield the Nash equilibrium
\[ u^*(x) \equiv 0 \] (4.6)
if \( \sup_{x \in [0,1]} \int_0^1 |r(x,y)|^2 \, dy < 1. \)

### 4.2 Convergence Results

To state our convergence results, we introduce two sets of functions. We define \( \Omega \) as the set of positive, bounded functions \( \omega : [0,1] \rightarrow \mathbb{R}_+ \) such that, at each element of the set \( \omega([0,1]) \cup 2\omega([0,1]) \), the level set of \( \omega \) is of measure zero. Let \( \Pi \) be the set of positive functions \( \nu(x) \) that are either strictly increasing or strictly decreasing. Then, \( \Pi \subset \Omega \). Also contained in \( \Omega \) are all bounded \( C^1[0,1] \) positive functions whose derivative is zero on a set of measure zero.

For games with payoff functions of the form (4.3), we have the following result:

**Theorem 4.1** Consider the system (4.1)–(4.2), along with (4.3) and (4.4), where \( k(x) = \varepsilon K(x) = O(\varepsilon) \), \( \varepsilon \) is a small, positive constant, and
\[ c(x) > \frac{1}{2} \left( \int_0^1 d^2(y) \, dy \right)^{1/2} \left( \int_0^1 p^2(y) \, dy \right)^{1/2} \]
for all $x \in [0, 1]$. There exists a constant $\varepsilon$ such that for all $\varepsilon \in (0, \varepsilon)$ and functions $\omega \in \Omega$, if the $L^2[0, 1]$ norm of $\Delta(x, 0)$ is sufficiently small, then for all $t \geq 0$,
\[
\int_0^1 \Delta^2(x, t) \, dx \leq M e^{-mt} \int_0^1 \Delta^2(x, 0) \, dx + O \left( \varepsilon^2 + \max_{x \in [0, 1]} a^2(x) \right),
\]
where
\[
\Delta(x, t) = u(x, t) - u^*(x),
\]
\[
M = \frac{\max_x \{k(x)a^2(x)\}}{\min_x \{k(x)a^2(x)\}},
\]
\[
m = 2 \min_{x \in [0, 1]} \{\alpha(x)\} \min_{x \in [0, 1]} \{k(x)a^2(x)\},
\]
\[
\alpha(x) = c(x) - \frac{1}{2} \left( \int_0^1 d^2(y) \, dy \right) \left( \int_0^1 p^2(y) \, dy \right) \frac{1}{2}.
\]

**Proof:** Denote the error at time $t$ relative to the Nash equilibrium as
\[
\tilde{u}(x, t) = u(x, t) - \mu(x, t) - u^*(x).
\]
By substituting (4.3) into (4.1)–(4.2), we obtain the error system
\[
\frac{\partial}{\partial t} \tilde{u}(x, t) = \varepsilon K(x) G[\tilde{u}, u^*, c, d, p, q, r, \varepsilon, \mu](x, t),
\]
where the operator $G$ (with the arguments suppressed) is defined as
\[
G[\cdot](x, t) \triangleq \mu(x, t) \left[ -c(x)(\tilde{u}(x, t) + u^*(x) + \mu(x, t))^2 
\right.
\]
\[
+ d(x)(\tilde{u}(x, t) + u^*(x) + \mu(x, t))
\]
\[
\times \int_0^1 p(y)(\tilde{u}(y, t) + u^*(y) + \mu(y, t)) \, dy
\]
\[
+ q(x)(\tilde{u}(x, t) + u^*(x) + \mu(x, t))
\]
\[
+ \int_0^1 r(x, y)(\tilde{u}(y, t) + u^*(y) + \mu(y, t)) \, dy + e(x) \right].
\]
The form of (4.13) admits the application of general averaging theory [45] for stability analysis, and the average error system can be shown to be
\[
\frac{\partial}{\partial t} \tilde{u}^{\text{ave}}(x, t) = \lim_{T \to \infty} \frac{\varepsilon K(x)}{T} \int_0^T G[\cdot](x, t) \, dt,
\]
\[
= -\varepsilon K(x)a^2(x) \left( c(x) \tilde{u}^{\text{ave}}(x, t) - \frac{1}{2} d(x) \int_0^1 p(y) \tilde{u}^{\text{ave}}(y, t) \, dy \right).
\]
Let \( V(t) \) be a Lyapunov functional defined as
\[
V(t) = \frac{1}{2\varepsilon} \int_0^1 \frac{1}{K(x)a^2(x)} (\tilde{u}^{\text{ave}})^2(x,t) \, dx
\] (4.16)
and bounded from both sides by
\[
V(t) \geq \frac{1}{2\varepsilon \max_x \{K(x)a^2(x)\}} \int_0^1 (\tilde{u}^{\text{ave}})^2(x,t) \, dx, \tag{4.17}
\]
\[
V(t) \leq \frac{1}{2\varepsilon \min_x \{K(x)a^2(x)\}} \int_0^1 (\tilde{u}^{\text{ave}})^2(x,t) \, dx. \tag{4.18}
\]

Taking the time derivative, substituting (4.15), and applying the Cauchy-Schwarz inequality yields
\[
\dot{V} = -\int_0^1 c(x) (\tilde{u}^{\text{ave}})^2(x,t) \, dx + \frac{1}{2} \int_0^1 d(x) \tilde{u}^{\text{ave}}(x,t) \, dx \int_0^1 p(y) \tilde{u}^{\text{ave}}(y,t) \, dy,
\]
\[
\leq -\int_0^1 c(x) (\tilde{u}^{\text{ave}})^2(x,t) \, dx + \frac{1}{2} \left( \int_0^1 |d(x)|^2 \, dx \right)^{\frac{1}{2}} \left( \int_0^1 |\tilde{u}^{\text{ave}}(x,t)|^2 \, dx \right)^{\frac{1}{2}}
\]
\[
\times \left( \int_0^1 |p(y)|^2 \, dy \right)^{\frac{1}{2}} \left( \int_0^1 |\tilde{u}^{\text{ave}}(y,t)|^2 \, dy \right)^{\frac{1}{2}}. \tag{4.19}
\]

Collecting terms and substituting the bound (4.18) gives
\[
\dot{V} \leq -\min_{x \in [0,1]} \{\alpha(x)\} \int_0^1 (\tilde{u}^{\text{ave}})^2(x,t) \, dx,
\]
\[
\leq -2\varepsilon \min_{x \in [0,1]} \{\alpha(x)\} \min_{x \in [0,1]} \{K(x)a^2(x)\} \cdot V. \tag{4.20}
\]

From the Comparison Lemma [55] and the bounds (4.17), (4.18), we obtain
\[
\int_0^1 (\tilde{u}^{\text{ave}})^2(x,t) \, dx \leq Me^{-mt} \int_0^1 (\tilde{u}^{\text{ave}})^2(x,0) \, dx. \tag{4.21}
\]

From [45, Theorem 3.6], the error system (4.13) retains the stability properties of the average system (4.15). Specifically,
\[
\int_0^1 \tilde{u}^2(x,t) \, dx \leq Me^{-mt} \int_0^1 \tilde{u}^2(x,0) \, dx + O(\varepsilon^2). \tag{4.22}
\]

Noting that \( u(x,t) - u^*(x) = \tilde{u}(x,t) + \mu(x,t) \) and \( \mu(x,t) \) is \( O(\max_x a(x)) \) completes the proof. \( \blacksquare \)
Similarly, for payoff functions of the form (4.5), we have the following result:

**Theorem 4.2** Consider the system (4.1)–(4.2), along with (4.5) and \( u^*(x) \equiv 0 \), where \( k(x) = \varepsilon K(x) = O(\varepsilon) \), \( \varepsilon \) is a small, positive constant, and

\[
\sup_{x \in [0,1]} \int_0^1 |r(x,y)|^2 \, dy \leq 1.
\]

There exists a constant \( \bar{\varepsilon} \) such that for all \( \varepsilon \in (0, \bar{\varepsilon}) \) and functions \( \omega \in \Omega \), if the \( L^2[0,1] \) norm of \( u(x,0) \) is sufficiently small, then for all \( t \geq 0 \),

\[
\int_0^1 u^2(x,t) \, dx \leq M e^{-\sigma t} \int_0^1 u^2(x,0) \, dx + O \left( \varepsilon^2 + \max_{x \in [0,1]} a^2(x) \right),
\]

where

\[
\sigma = 2 \beta \min_{x \in [0,1]} \{ k(x)a^2(x) \},
\]

\[
\beta = 1 - \sup_{x \in [0,1]} \left( \int_0^1 |r(x,y)|^2 \, dy \right)^{1/2},
\]

and \( M \) is given by (4.9).

**Proof:** Following the proof of Theorem 4.1, we obtain the average error system,

\[
\frac{\partial}{\partial t} \bar{u}^{\text{ave}}(x,t) = -\varepsilon K(x)a^2(x) \left( \bar{u}^{\text{ave}}(x,t) - \int_0^1 r(x,y)\bar{u}^{\text{ave}}(y,t) \, dy \right),
\]

and prove its exponential stability using the Lyapunov functional (4.16). Specifically, we have

\[
\dot{V} = -\int_0^1 (\bar{u}^{\text{ave}})^2(x,t) \, dx + \int_0^1 \bar{u}^{\text{ave}}(x,t) \int_0^1 r(x,y)\bar{u}^{\text{ave}}(y,t) \, dy \, dx,
\]

\[
\leq -\int_0^1 (\bar{u}^{\text{ave}})^2(x,t) \, dx + \sup_{x \in [0,1]} \left( \int_0^1 |r(x,y)|^2 \, dy \right)^{1/2}
\times \left( \int_0^1 |\bar{u}^{\text{ave}}(x,t)|^2 \, dx \right)^{1/2} \left( \int_0^1 |\bar{u}^{\text{ave}}(y,t)|^2 \, dy \right)^{1/2},
\]

\[
\leq -2\varepsilon \beta \min_{x \in [0,1]} \{ K(x)a^2(x) \} V,
\]

where we have applied the Cauchy-Schwarz inequality and substituted the bound (4.18). The remainder of the proof follows directly after noting \( u^*(x) \equiv 0 \).
The structural differences between these two classes of quadratic payoff functions are manifested in their respective convergence rates—specifically, in the terms (4.11) and (4.25). To understand more intuitively how these convergence results are achieved, one may also compute the average of the \( \hat{u}(x,t) \)-system (4.1), which reveals that this Nash seeking method is, on average, a gradient descent algorithm.

### 4.3 Uncountably-Many Player Price Game

For an example game with uncountably-many players, we consider a price game where the players, indexed by \( x \in [0,1] \), wish to maximize payoff functions of the form

\[
J(x,t) = s(x,t) (u(x,t) - m(x)),
\]

where \( s(x,t) \) is the sales volume of player \( x \), \( u(x,t) \) is its price, and \( m(x) \) is its marginal cost.

The sales volume \( s(x,t) \) is modeled as

\[
s(x,t) = \frac{R_\parallel}{R(x)} \left( S - \frac{u(x,t)}{R_\parallel} + \int_0^1 \frac{u(y,t)}{R(y)} dy \right),
\]

where \( S \) is the total sales volume and \( R(x) \) is the consumer’s “resistance” toward buying the product of player \( x \). This resistance \( R(x) \) is measurable, positive, and bounded for all \( x \in [0,1] \). We assume the players do not know the mathematical structure of (4.28) as it is difficult to know the consumer preference and how it enters the sales model (4.29).

Substituting (4.29) into (4.28) leads to

\[
J(x,t) = -\frac{u^2(x,t)}{R(x)} + \frac{R_\parallel}{R(x)} u(x,t) \int_0^1 \frac{u(y,t)}{R(y)} dy + \left( \frac{R_\parallel S}{R(x)} + \frac{m(x)}{R(x)} \right) u(x,t)
\]

\[
- \frac{R_\parallel m(x)}{R(x)} \int_0^1 \frac{u(y,t)}{R(y)} dy - \frac{R_\parallel Sm(x)}{R(x)},
\]

(4.31)
which yields the Nash equilibrium prices,

\[ u^*(x) = R_{||} \left( S + \frac{1}{2} \frac{m(x)}{R_{||}} + \frac{1}{2} \int_0^1 \frac{m(y)}{R(y)} \, dy \right). \]  

(4.32)

Because (4.31) has the form of (4.3), the conditions of Theorem 4.1 are satisfied when the players implement the Nash seeking strategy (4.1)–(4.2). Hence, the players achieve local, exponential convergence to Nash equilibrium \( u^*(x) \). For this specific example, one can obtain

\[ \int_0^1 \Delta^2(x,t) \, dx \leq Me^{-\xi t} \int_0^1 \Delta^2(x,0) \, dx + O \left( \varepsilon^2 + \max_{x \in [0,1]} a^2(x) \right), \]  

(4.33)

where \( \Delta, M \) are given by (4.8), (4.9). The convergence rate

\[ \xi = \frac{\min_x \{k(x)a^2(x)\}}{\max_x \{R(x)\}} \]

is found by performing Lyapunov analysis on the average error system using the Lyapunov functional (4.16).

For a numerical example, we choose the parameters \( S = 100, m(x) = 20 + 5 \sin(4\pi x) \), and \( R(x) = 1 + \cos(2\pi x)/4 \), which result in the Nash equilibrium

\[ u^*(x) = 25\sqrt{15} + 20 + \frac{5}{2} \sin(4\pi x), \]  

(4.34)

and the corresponding sales volume

\[ s^*(x) = \frac{100\sqrt{15} - 10 \sin(4\pi x)}{4 + \cos(2\pi x)}. \]  

(4.35)
Figure 4.3: Nash equilibrium $u^*(x)$ (dashed) of the oligopoly price game with the players’ price (blue) at $t = 400$ sec superimposed.

To approximate the uncountably-many player game, we discretize the interval $[0, 1]$ using $N = 1001$ points, representing $N$ players, and use the trapezoidal numerical integration scheme to approximate the integral terms in (4.31). For these $N$ players, we select the Nash seeking parameters $k(x)$, $a(x)$, and $\omega(x)$ by a random draw from a uniform distribution. Namely, $k(x)$ is sampled from the distribution $U^k(1, 5)$, $a(x)$ from $U^a(0.1, 0.2)$, and $\omega(x)$ from $U^\omega(30, 60)$, where $U(a, b)$ denotes the uniform distribution with probability density function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise}. \end{cases} \quad (4.36)$$

The players are initialized at $u(x, 0) = S + v(x)$ where $v$ is distributed normally with zero-mean and unity variance.

Figure 4.2 depicts the evolution of the players’ prices as they converge to a neighborhood of $u^*(x)$. Figure 4.3 shows $u^*(x)$ with $\hat{u}(x, t)$ at $t = 400$ sec. We show $\hat{u}(x, t)$ to highlight the players’ convergence to $u^*(x)$ since $u(x, t)$ contains the additive signal $\mu(x, t)$. Figure 4.4 depicts the evolution of the players’ payoff values.
4.4 Summary

We have introduced a non-model based approach to solve noncooperative games with uncountably-many players that possess quadratic payoff functions. A player can stably attain its Nash equilibrium by measuring only the value of its payoff function. No other information about the game is needed. One potential application of this approach may be real-time price negotiation in large electronic markets as the supply and demand fluctuates.


The dissertation author is the primary investigator and author of this work.
Chapter 5

Finite-Horizon LQ Control for Unknown Discrete-Time Systems via Extremum Seeking

We solve here a finite-time horizon, optimal control problem for a completely unknown discrete-time linear system by using an extremum seeking controller to attain the open-loop optimal control sequence. Since only the output values of the system are known, an optimal feedback control is not possible and the considered cost functions are restricted to quadratic functions in the input and the output, rather than the state. The control values for each step in the time horizon are determined by the extremum seeking controller, which is driven by the value of the cost multiplied by sinusoidal perturbations. After each iteration of the unknown system, the cost is computed and used to iterate the extremum seeking controller, which updates the control sequence. On average, the controller is driven by the gradient of the cost function, which is driven to zero, causing the control sequence to converge to the optimal control.

Our results follow closely and extend those found in [23] to establish convergence for multi-variable, discrete-time extremum seeking controllers. Convergence of multi-variable, discrete-time extremum seeking schemes with noisy measurements has been shown in [100], but the proof relies on diminishing gains that effectively turn off the algorithm over time. We utilize the two-time scale av-
eraging theory for discrete time systems [5] to prove convergence for a multi-
parameter scheme with constant gains, albeit with noise-free measurements, so
that the scheme continually seeks to optimize the measured cost, tracking changes
in the system. While we prove a local result, non-local results for extremum seeking
controllers have been established [103].

Systems that are amenable to this framework, i.e., unknown or uncertain
systems that are highly repeatable with constant initial conditions, have been the
focus of a large body of research known as iterative learning control [22, 72, 109].
Pulsed laser systems, for example, may fit the above description. For unknown
systems, iterative learning control utilizes finite-time, repetitive experiments to
achieve the tracking of a given reference, whereas in this work, our scheme op-
timizes the system’s output according to the computed cost function. Stability
of the unknown system is not required (since we are dealing with a finite-horizon
problem); however, the optimization problem becomes more difficult as the level of
instability of the unknown system increases or as the time horizon increases for a
scalar system. We assume that the unknown system is reachable to ensure that an
optimal control exists to drive the system from any initial state to any final state if
there are no hard input constraints. If the system is subject to input constraints,
our approach can accommodate them by utilizing a projection operator, as shown
in Section 5.5.

This chapter is organized as follows: We introduce the general problem
statement in Section 5.1, discuss the extremum seeking controller in Section 5.2,
and state our convergence result in Section 5.4. We provide simulation results in
Section 5.5 and conclude with Section 5.6.

5.1 Problem Statement

Consider the single-input linear discrete-time system

\[ x_{k+1} = A_k x_k + B_k u_k, \]  
\[ y_k = C_k x_k, \]
where \( x_k \in \mathbb{R}^n \), \( u_k \in \mathbb{R} \), \( y_k \in \mathbb{R}^p \), and \( A_k, B_k, C_k \) denote unknown matrices of appropriate dimensions at discrete time \( k \). Only the input \( u_k \) and the output \( y_k \) are treated as known values. Hence, even the state dimension is unknown.

Despite these significant uncertainties, we wish to find the optimal, open-loop control sequence \( \{u^*_k\}_{k=0}^{N-1} \) that minimizes the cost function

\[
J(u) = \frac{1}{2} y_N^T \hat{Q}_N y_N + \sum_{k=0}^{N-1} y_k^T \hat{Q}_k y_k + R_k u_k^2,
\]

where \( u = [u_0, \ldots, u_{N-1}]^T \) and \( \hat{Q}_k, \hat{Q}_N \geq 0 \), \( R_k > 0 \) for all \( k \in \{0, \ldots, N-1\} \).

Namely, we want to solve the discrete-time, finite-time horizon optimal control problem,

\[
\min_u J(u), \text{ subject to } (5.1)-(5.2)
\]

with initial condition \( x_0 \). We seek an open-loop solution instead of an optimal state feedback policy since the system is unknown and state information is not available. The cost function (5.3) can be written in terms of the state \( x_k \) as

\[
J(u) = \frac{1}{2} x_N^T Q_N x_N + \sum_{k=0}^{N-1} x_k^T Q_k x_k + R_k u_k^2,
\]

where the state penalty matrices are

\[
Q_k = C_k^T \hat{Q}_k C_k, \quad k \in \{0, \ldots, N\},
\]

which is the standard cost function for the LQ optimal control problem with a positive weight on each control stage, and thus, a unique minimizing control policy exists.

By substituting the system’s state trajectory,

\[
x_k = \Phi_{k,0} x_0 + \sum_{l=0}^{k-1} \Phi_{k,l+1} B_l u_l,
\]

where

\[
\Phi_{i,j} = \begin{cases} 
A_{i-1}A_{i-2} \cdots A_j, & i > j \\
I, & i = j
\end{cases}
\]

(5.8)
the cost function (5.5) can be written in terms of only the initial state $x_0$ and the control $u_k$. In vector form, we have

$$J(u) = \frac{1}{2} x_0^T F x_0 + x_0^T G u + \frac{1}{2} u^T H u,$$

(5.9)

where

$$F = Q_0 + \Theta^T \bar{Q} \Theta,$$

(5.10)

$$G = \Theta^T Q \Lambda,$$

(5.11)

$$H = \Lambda^T \bar{Q} \Lambda + \bar{R},$$

(5.12)

$$\Theta = [\Phi_{1,0}, \Phi_{2,0}, \ldots, \Phi_{N,0}]^T,$$

(5.13)

$$\bar{Q} = \text{diag}[Q_1, \ldots, Q_N],$$

(5.14)

$$\Lambda = \begin{bmatrix}
B_0 & 0 & 0 & \cdots & 0 \\
\Phi_{2,1} B_0 & B_1 & 0 & \cdots \\
\Phi_{3,1} B_0 & \Phi_{3,2} B_1 & B_2 & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
\Phi_{N,1} B_0 & \Phi_{N,2} B_1 & \Phi_{N,3} B_2 & B_{N-1}
\end{bmatrix},$$

(5.15)

$$\bar{R} = \text{diag}[R_0, \ldots, R_{N-1}],$$

(5.16)

and diag[…] denotes a block diagonal matrix. The Hessian $H$ is positive definite symmetric since it equals the sum of $\Lambda^T \bar{Q} \Lambda \succeq 0$ and $\bar{R} > 0$. If the system model were known, the optimal open-loop control values could be found directly by solving the corresponding quadratic program [26, 17].

The eigenvalues of $H$, denoted by $\lambda(H)$, play a prominent role in the convergence rate of our algorithm, and below, we provide bounds on their values. These bounds are available because $H$ equals the sum of a positive semidefinite symmetric matrix and a positive definite symmetric matrix—a topic that has been the focus of much research [15, 41].

**Lemma 5.1** Let $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_N$ be the ordered eigenvalues of $H$. Similarly, let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ and $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_N$ be the ordered eigenvalues of $\Lambda^T \bar{Q} \Lambda$ and $\bar{R}$, respectively. Then,

$$\max_{i+j=N+l} \{\lambda_i + \rho_j\} \leq \beta_l \leq \min_{i+j=l+1} \{\lambda_i + \rho_j\}.$$  

(5.17)
If the input penalty is constant, i.e., $R_k \equiv R$, then

$$\beta_l = \lambda_l + R. \quad (5.18)$$

**Proof:** Since $H = \Lambda^T \bar{Q} \Lambda + \bar{R}$, where $\Lambda^T \bar{Q} \Lambda \geq 0$ and $\bar{R} > 0$, (5.17) is immediate [15, 41]. If the input penalty is constant, $\rho_j = R$ for all $j$, so

$$\max_{i+j=N+l}\{\lambda_i + \rho_j\} = \min_{i+j=l+1}\{\lambda_i + \rho_j\} = \lambda_l + R,$$

resulting in (5.18).

It is important to note that $\lambda(H)$ is bounded below by the smallest input penalty weight $\min_k R_k$, which is a strictly positive value. Since our system is unknown and possibly high dimensional, the matrix $\Lambda^T \bar{Q} \Lambda$ is likely to have some, if not many, zero eigenvalues due to the sparsity of $\bar{Q}$, meaning that a subset of $\{\beta_l\}_{l=1}^N$ will be precisely equal to $R$ when the input penalty is constant.

We now make the following assumption for (5.1):

**Assumption 5.1** The reachability Gramian [87],

$$W_N = \sum_{l=0}^{N-1} \Phi_{N,l+1} B_l B_l^T \Phi_{N,l+1}^T, \quad (5.19)$$

is positive definite symmetric.

This assumption is a necessary and sufficient condition for the existence of a control sequence to take the system from any initial state $x_0$ to any final state $x_N$, provided there are no hard constraints on the control.

### 5.2 Optimal Control via Extremum Seeking

To attain the optimal open-loop control sequence for the unknown system (5.1)–(5.2) with the cost function (5.3), we employ the non-model based optimization strategy known as extremum seeking, which uses sinusoidal perturbations to estimate, and drive to zero, the gradient of an unknown function, yielding the function’s minimizer. In this case, the unknown function is the cost function (5.9) and the minimizer is $\{u^*_k\}_{k=0}^{N-1}$.

To implement this optimization scheme, we drive the system (5.1)–(5.2) with the control sequence $\{u_k^*\}_{k=0}^{N-1}$ to obtain the cost value $J(l)$ (5.3), where $l$
First, we rewrite the cost function (5.9) using its Taylor series expansion about the $k$-th iteration of the algorithm. The discrete-time extremum seeking controller is then iterated with $J(l)$ as the input to produce an updated control sequence $\{u_k^{l+1}\}_{k=0}^{N-1}$. The extremum seeking controller is given by

$$\hat{u}_k(l) = -\frac{\gamma_k}{z - 1}[\xi_k(l)],$$

$$\xi_k(l) = \cos(\omega_k l - \varphi_k) \frac{z - 1}{z + h} J(l),$$

$$u_k(l) = \hat{u}_k(l) + a_k \cos(\omega_k l),$$

for $k \in \{0, \ldots, N-1\}$, where $\gamma_k = \epsilon \Gamma_k > 0$, $\epsilon$ is a small, positive parameter, $h \in (0, 1)$, $a_k > 0$, and $l$ denotes the discrete time index of the extremum seeking controller. The selected modulation frequencies are given by $\omega_k = b_k \pi$, where $|b_k| \in (0, 1)$ is a rational number and $\omega_k \neq \omega_i$ for all distinct $i, k \in \{0, \ldots, N-1\}$.

The notation $P(z)[q(l)]$ is used to denote the time-domain signal that is the output of the transfer function $P(z)$ driven by $q(l)$.

### 5.3 Error System

To analyze the convergence of the system, we study its error dynamics. First, we rewrite the cost function (5.9) using its Taylor series expansion about the
optimum $u^* = [u_0^*, \ldots, u_{N-1}^*]^T$, 

$$J(u) = J(u^*) + \frac{1}{2}(u - u^*)^T H(u - u^*), \tag{5.23}$$

where the Hessian $H$ is constant since the cost function is quadratic, and we note that $\nabla J(u^*)$, the gradient of $J$ at $u^*$, is zero. Next, denote the error relative to the optimal control sequence as

$$\tilde{u}_k(l) = u_k(l) - a_k \cos(\omega_k l) - u_k^*,$$

$$= \tilde{u}_k(l) - u_k^*, \tag{5.24}$$

and substitute into (5.23) to obtain

$$J(\tilde{u}) = J(u^*) + \frac{1}{2}(\tilde{u} + \mu)^T H(\tilde{u} + \mu)$$

$$= J(u^*) + \frac{1}{2} \tilde{u}^T H \tilde{u} + \mu^T H \tilde{u} + \frac{1}{2} \mu^T H \mu, \tag{5.25}$$

where $\mu = [a_0 \cos(\omega_0 l), \ldots, a_{N-1} \cos(\omega_{N-1} l)]^T$ and $\tilde{u} = [\tilde{u}_0, \ldots, \tilde{u}_{N-1}]^T$.

Substituting (5.24) into (5.20) yields

$$\tilde{u}_k(l) = -\frac{\epsilon \Gamma_k}{z - 1} \xi_k(l) - u_k^*, \tag{5.26}$$

which, after noting (5.21) and (5.25), can be expressed as the difference equation

$$\tilde{u}_k(l + 1) = \tilde{u}_k(l) - \epsilon \Gamma_k c_k(l) W(z)[J(u^*)] - \frac{1}{2} \epsilon \Gamma_k c_k(l) W(z)[\tilde{u}^T H \tilde{u} + \mu^T H \mu]$$

$$- \epsilon \Gamma_k c_k(l) W(z)[\mu^T H \tilde{u}], \tag{5.27}$$

for all $k \in \{0, \ldots, N - 1\}$, where $c_k(l) = \cos(\omega_k l - \varphi_k)$ and $W(z)$ denotes the washout filter $(z - 1)/(z + h)$. Each difference equation is driven by the same value $W(z)[J(l)]$ but with sinusoidal perturbations of different frequencies.

Applying Lemmas B.1, B.2, and B.3 from Appendix B to the last term of (5.27) yields, after some algebra,

$$c_k(l) W(z)[\mu^T H \tilde{u}] = \frac{1}{2} a_k \text{Re}\{e^{j\varphi_k} W(e^{j\omega_k z})[h_k^T \tilde{u}]\}$$

$$+ \frac{1}{2} c_k^-(l) \text{Re}\{\Omega(z)[H \tilde{u}]\} - \frac{1}{2} S_k^-(l) \text{Im}\{\Omega(z)[H \tilde{u}]\}$$

$$+ \frac{1}{2} c_k^+(l) \text{Re}\{\Omega(z)[H \tilde{u}]\} - \frac{1}{2} S_k^+(l) \text{Im}\{\Omega(z)[H \tilde{u}]\}, \tag{5.28}$$
where \( h_k \) denotes the \( k \)-th column of \( H \),

\[
C_k^-(l)^T \triangleq \begin{bmatrix}
a_0 \cos((\omega_0 - \omega_k)l + \varphi_k) \\
\vdots \\
a_{k-1} \cos((\omega_{k-1} - \omega_k)l + \varphi_k) \\
0 \\
a_{k+1} \cos((\omega_{k+1} - \omega_k)l + \varphi_k) \\
\vdots \\
a_{N-1} \cos((\omega_{N-1} - \omega_k)l + \varphi_k)
\end{bmatrix}
\tag{5.29}
\]

\[
S_k^-(l)^T \triangleq \begin{bmatrix}
a_0 \sin((\omega_0 - \omega_k)l + \varphi_k) \\
\vdots \\
a_{k-1} \sin((\omega_{k-1} - \omega_k)l + \varphi_k) \\
0 \\
a_{k+1} \sin((\omega_{k+1} - \omega_k)l + \varphi_k) \\
\vdots \\
a_{N-1} \sin((\omega_{N-1} - \omega_k)l + \varphi_k)
\end{bmatrix}
\tag{5.30}
\]

\[
C_k^+(l)^T \triangleq \begin{bmatrix}
a_0 \cos((\omega_0 + \omega_k)l - \varphi_k) \\
\vdots \\
a_{N-1} \cos((\omega_{N-1} + \omega_k)l - \varphi_k)
\end{bmatrix}
\tag{5.31}
\]

\[
S_k^+(l)^T \triangleq \begin{bmatrix}
a_0 \sin((\omega_0 + \omega_k)l - \varphi_k) \\
\vdots \\
a_{N-1} \sin((\omega_{N-1} + \omega_k)l - \varphi_k)
\end{bmatrix}
\tag{5.32}
\]

\[
\Omega(z) \triangleq \text{diag}[W(e^{j\omega_1 z}), \ldots, W(e^{j\omega_{N-1} z})].
\tag{5.33}
\]

Substituting (5.28) into (5.27) allows the error dynamics to be expressed as follows:

\[
\ddot{u}_k(l+1) - 2\delta_k(l) = \epsilon(L_k(z)[h_k^T \ddot{u}] + \Psi_{k,1}^-(l) + \Psi_{k,1}^+(l) + \Psi_{k,2}(l)) + \delta_k(l),
\tag{5.34}
\]

where

\[
L_k(z) = -\frac{1}{4} \Gamma_k a_k(e^{j\varphi_k}W(e^{j\omega_k z}) + e^{-j\varphi_k}W(e^{-j\omega_k z}))
\tag{5.35}
\]

\[
\Psi_{k,1}^-(l) = \frac{1}{2} \Gamma_k S_k^-(l) \text{Im}\{\Omega(z)[H \ddot{u}]\} - \frac{1}{2} \Gamma_k C_k^-(l) \text{Re}\{\Omega(z)[H \ddot{u}]\},
\tag{5.36}
\]
\[
\Psi_{k,1}^+(l) = \frac{1}{2} \Gamma_k S_k^+(l) \text{Im}\{\Omega(z)[H\tilde{u}]\} - \frac{1}{2} \Gamma_k C_k^+(l) \text{Re}\{\Omega(z)[H\tilde{u}]\},
\]
(5.37)
\[
\Psi_{k,2}(l) = -\Gamma_k c_k(l) W(z)[\tilde{u}^T H\tilde{u}],
\]
(5.38)
\[
\delta_k(l) = -\epsilon \Gamma_k c_k(l) W(z)[J(u^*) + \mu^T H\mu].
\]
(5.39)

In the above formulation, one sees that the error dynamics evolve according to a sum of a linear time-invariant term, \( L_k(z)[h_k^T \tilde{u}] \); linear time-varying function, \( \Psi_{k,1}(l) \) and \( \Psi_{k,1}^+(l) \); a nonlinear time-varying function \( \Psi_{k,2}(l) \); and a time-varying function \( \delta_k(l) \) that does not depend on the control error \( \tilde{u} \).

The following lemma states the convergence properties of \( \delta_k(l) \):

**Lemma 5.2** The time-varying function \( \delta_k(l) \) exponentially converges to an \( O(\epsilon \bar{a}^2) \)-neighborhood of zero, where \( \bar{a}^2 = \max_i \{a_i^2\} \).

**Proof:** Rewrite (5.39) as \( \delta_k(l) = \delta_{1,k}(l) + \delta_{2,k}(l) \), where
\[
\delta_{1,k}(l) = -\epsilon \Gamma_k c_k(l) W(z)[J(u^*)],
\]
(5.40)
\[
\delta_{2,k}(l) = -\epsilon \Gamma_k c_k(l) W(z)[\mu^T H\mu].
\]
(5.41)

Since the washout filter \( W(z) \) has zero DC gain, \( \delta_{1,k}(l) \) has only exponentially decaying terms. For \( \delta_{2,k}(l) \), we rewrite the argument of \( W(z) \) as a double summation to obtain
\[
\delta_{2,k}(l) = -\epsilon \Gamma_k c_k(l) W(z) \left[ \sum_{i=1}^{N} \sum_{j=1}^{N} h_{ij} \mu_i \mu_j \right],
\]
(5.42)
where \( h_{ij} \) denotes the \( i, j \)-th element of \( H \). After applying a trigonometric identity to split the product \( \mu_i \mu_j \) into two terms and invoking the linearity of the washout filter, generating a summation of \( 2N^2 \) filtered signals, we obtain the following conservative bound:
\[
\delta_{2,k}(l) \leq \epsilon \chi \max_i \{a_i^2\},
\]
(5.43)
where \( \chi = 2N^2 \Gamma_k \lambda_{\max}(H) \Xi \), \( \lambda_{\max}(H) \) is the largest eigenvalue of \( H \), and \( \Xi \) is the maximum value of \( |W(e^{j\sigma})| \) in the summation, with \( \sigma \) denoting the frequencies obtained after applying the trigonometric identity.
In the next section, we state our main result and prove convergence of $u$ to a neighborhood of the optimum $u^*$. First, we prove local exponential convergence of the error system (5.34) without the $\delta_k(l)$ by employing the two-time scale averaging theory for discrete-time systems [5] before considering the full system with $\delta_k(l)$.

### 5.4 Convergence Result

To establish the convergence result for the multi-variable, discrete-time extremum seeking scheme, we extend the results found in [23]. The following theorem states a sufficient condition for locally exponential convergence of the error system (5.34) without the $\delta_k(l)$ term. Namely, for all $k \in \{0, \ldots, N-1\}$, we have the homogeneous error system that is periodic in time $l$,

$$
\tilde{u}_k(l + 1) - \tilde{u}_k(l) = \epsilon (L_k(z)[h_k^T \tilde{u}] + \Psi_{k,1}(l) + \Psi_{k,1}^+(l) + \Psi_{k,2}(l)).
$$

(5.44)

**Theorem 5.1** Consider the homogeneous error system (5.44) with modulation frequencies that satisfy $\omega_k \neq \omega_i$ for all distinct $i, k \in \{0, \ldots, N-1\}$. There exists a positive constant $\epsilon^*$ such that for all $\epsilon \in (0, \epsilon^*)$, the state-space realization of (5.44) is locally exponentially stable at the origin.

**Proof:** Since the error dynamics (5.44) depend on both linear and quadratic terms of $\tilde{u}$, we utilize the two-time scale averaging theory [5]. We note that due to the exponential stability of the washout filter $W(z)$, minimal state space realizations of $L_k(z), \Omega(z)$, and $W(z)$ can be chosen as $(\bar{A}_1, \bar{B}_1, \bar{C}_1, \bar{D}_1)$, $(\bar{A}_2, \bar{B}_2, \bar{C}_2, \bar{D}_2)$, and $(\bar{A}_3, \bar{B}_3, \bar{C}_3, \bar{D}_3)$, where $\bar{A}_1$, $\bar{A}_2$, and $\bar{A}_3$ are exponentially stable, i.e., their poles are inside the unit circle. These realizations allow (5.44) to be written in state-space form as

$$
\zeta'(l + 1) = \bar{A}_1 \zeta'(l) + h(l, \tilde{u}(l)),
$$

(5.45)

$$
\tilde{u}_k(l + 1) = \tilde{u}_k(l) + \epsilon f'(l, \tilde{u}_k(l), \zeta'(l)).
$$

(5.46)
where

\[ \bar{A} = \text{diag}[\bar{A}_1, \bar{A}_2, \bar{A}_3], \]  
\[ h(l, \bar{u}(l)) = [\bar{u}^T h_k \bar{B}_1^T | \bar{u}^T H \bar{B}_2^T | \bar{u}^T H \bar{u} \bar{B}_3^T]^T, \]  
\[ f'(l, \bar{u}(l), \zeta'(l)) = \bar{D}_1 h_k^T \bar{u} + \eta_k(l) \bar{D}_2 H \bar{u} + c_k(l) \bar{D}_3 \bar{u}^T H \bar{u} \]  
\[ + [\bar{C}_1 | \eta_k(l) \bar{C}_2 | c_k(l) \bar{C}_3] \zeta'(l), \]  
\[ \eta_k(l) = \eta_k^-(l) + \eta_k^+(l), \]  
and the functions \( \eta_k^-(l) \) and \( \eta_k^+(l) \) depend on \( C_k^- \) and \( C_k^+ \), respectively, such that \( \eta_k^-(l)([0 | \bar{C}_2 | 0] x'(l) + \bar{D}_2 H \bar{u}) = \Psi_{k,1}^- \) and \( \eta_k^+(l)([0 | \bar{C}_2 | 0] \zeta'(l) + \bar{D}_2 H \bar{u}) = \Psi_{k,1}^+ \). The mixed time scale system (5.45)–(5.46), with the exponentially stable matrix \( \bar{A} \), can be transformed to admit the application of the two-time scale averaging theory [5].

Let

\[ \zeta(l) = \zeta'(l) - w(l, \bar{u}), \]  
\[ w(l, \bar{u}) = \sum_{i=0}^{l-1} \bar{A}^{l-i-1} h(i, \bar{u}), \]

with \( h(l, \bar{u}) \) shown in (5.48). The transformed system is then

\[ \zeta(l + 1) = \bar{A} \zeta(l) + \epsilon g(l, \bar{u}, \zeta), \]  
\[ \bar{u}(l + 1) = \bar{u}(l) + \epsilon f(l, \bar{u}, \zeta), \]

where

\[ \epsilon g(l, \bar{u}, \zeta) = \bar{A} w(l, \bar{u}(l)) + h(l, \bar{u}(l)) - w(l + 1, \bar{u}(l + 1)), \]  
\[ = w(l + 1, \bar{u}(l)) - w(l + 1, \bar{u}(l + 1)), \]  
\[ = -\epsilon f'(l, \bar{u}, \zeta + w(l, \bar{u})) \int_0^1 \frac{\partial w}{\partial \bar{u}}(l + 1, s \bar{u}(l + 1) + (1 - s) \bar{u}(l)) \, ds, \]  
\[ f(l, \bar{u}, \zeta) = f'(l, \bar{u}(l), \zeta + w(k, \bar{u})). \]

To prove convergence, we analyze the averaged system of (5.53), given by

\[ \bar{u}_k^{ave}(l + 1) = \bar{u}_k^{ave}(l) + \epsilon f^{ave}(\bar{u}_k^{ave}(l)), \]
where \( f^{\text{ave}} \) is computed using the averaging operator \( \mathcal{A}\{\cdot\} \) [5],

\[
f^{\text{ave}}(\tilde{u}) = \mathcal{A}\{f(l, \tilde{u}, 0)\} = \lim_{T \to \infty} \frac{1}{T} \sum_{l=s+1}^{s+T} f(l, \tilde{u}, 0). \tag{5.57}
\]

Rather than attempting to compute \( f^{\text{ave}}(\tilde{u}) \) directly from (5.57), we use (5.49), (5.50), (5.51), and (5.55) to obtain

\[
f(l, \tilde{u}, 0) = f'(l, \tilde{u}, w(l, \tilde{u})),
\]

\[
= \tilde{D}_1 h_k^T \tilde{u} + \eta_k(l) \tilde{D}_2 H \tilde{u} + c_k(l) \tilde{D}_3 \tilde{u}^T H \tilde{u}
\]

\[
\quad + [\tilde{C}_1 \mid \eta_k(l) \tilde{C}_2 \mid c_k(l) \tilde{C}_3] \sum_{i=0}^{l-1} A^{l-i-1} [\tilde{u}^T h_k \bar{B}_1^T \mid \tilde{u}^T H \bar{B}_2^T \mid \tilde{u}^T H \bar{B}_3]^T,
\]

\[
= L_k(z)[h_k^T \tilde{u}] + \Psi_{k,1}^-(l) + \Psi_{k,1}^+(l) + \Psi_{k,2}(l), \tag{5.58}
\]

where \( \tilde{u} \) is considered a constant, and thus,

\[
f^{\text{ave}}(\tilde{u}) = \mathcal{A}\{L_k(z)[h_k^T \tilde{u}(l)] + \Psi_{k,1}^-(l) + \Psi_{k,1}^+(l) + \Psi_{k,2}(l)\}. \tag{5.59}
\]

Computing the individual terms yields

\[
\mathcal{A}\{\Psi_{k,1}^-(l)\} = \mathcal{A}\left\{\frac{1}{2} \Gamma_k S_k^-(l) \text{Im}\{\Omega(z)[H \tilde{u}(l)]\} - \frac{1}{2} \Gamma_k C_k^-(l) \text{Re}\{\Omega(z)[H \tilde{u}(l)]\}\right\},
\]

\[
= 0, \tag{5.60}
\]

\[
\mathcal{A}\{\Psi_{k,1}^+(l)\} = \mathcal{A}\left\{\frac{1}{2} \Gamma_k S_k^+(l) \text{Im}\{\Omega(z)[H \tilde{u}(l)]\} - \frac{1}{2} \Gamma_k C_k^+(l) \text{Re}\{\Omega(z)[H \tilde{u}(l)]\}\right\},
\]

\[
= 0, \tag{5.61}
\]

\[
\mathcal{A}\{\Psi_{k,2}(l)\} = \mathcal{A}\left\{-\Gamma_k c_k(l) W(z)[\tilde{u}(l)^T H \tilde{u}(l)]\right\},
\]

\[
= 0, \tag{5.62}
\]

where, again, \( \tilde{u}(l) \) is considered a constant sequence and we have noted that \( \omega_k \neq \omega_i \) for all \( i \neq k \). The final term, which yields the averaged system, is

\[
f^{\text{ave}}(\tilde{u}) = \mathcal{A}\{L_k(z)[h_k^T \tilde{u}(l)]\},
\]

\[
= \mathcal{A}\{-\frac{1}{4} \Gamma_k a_k [e^{i\omega_k} W(e^{i\omega_k} z) + e^{-i\omega_k} W(e^{-i\omega_k} z)] [h_k^T \tilde{u}(l)]\},
\]

\[
= -\rho_k \Gamma_k a_k h_k^T \tilde{u}, \tag{5.63}
\]
Lemma 5.2, \vspace{1em}

For all \( W(e^{i\omega k}) \) and \( \psi_k = \frac{\angle W(e^{i\omega k})}{\|W(e^{i\omega k})\|} \), the averaged system is

\[
\tilde{u}_k^{\text{ave}}(l + 1) = \tilde{u}_k^{\text{ave}}(l) - \epsilon \rho_k \Gamma_k a_k (h_k^T \tilde{u}_k^{\text{ave}}(l)),
\]

(5.64)

which depends on \( \tilde{u}^{\text{ave}} \) and in vector form is

\[
\tilde{u}^{\text{ave}}(l + 1) = (I - \epsilon KH) \tilde{u}^{\text{ave}}(l),
\]

(5.65)

where \( I \) is the \( N \times N \) identity matrix and \( K = \text{diag}[\rho_0 \Gamma_0 a_0, \ldots, \rho_{N-1} \Gamma_{N-1} a_{N-1}] \).

Let \( V(\tilde{u}^{\text{ave}}) = (\tilde{u}^{\text{ave}})^T K^{-1} \tilde{u}^{\text{ave}} \) be a Lyapunov function. Then, computing the difference \( \Delta V(\tilde{u}^{\text{ave}}) = V(\tilde{u}^{\text{ave}}(l + 1)) - V(\tilde{u}^{\text{ave}}(l)) \) yields

\[
\Delta V(\tilde{u}^{\text{ave}}) = (\tilde{u}^{\text{ave}})^T (I - \epsilon KH)^T K^{-1} (I - \epsilon KH) \tilde{u}^{\text{ave}} - (\tilde{u}^{\text{ave}})^T K^{-1} \tilde{u}^{\text{ave}},
\]

\[
= (\tilde{u}^{\text{ave}})^T (-2 \epsilon H + \epsilon^2 HKH) \tilde{u}^{\text{ave}},
\]

\[
\leq -2 \epsilon \lambda_{\text{min}}(H) |\tilde{u}^{\text{ave}}|^2 + \epsilon^2 \lambda_{\text{max}}(HKH) |\tilde{u}^{\text{ave}}|^2,
\]

(5.66)

where \( |\cdot| \) denotes the Euclidean norm. Selecting \( \epsilon = \frac{\lambda_{\text{min}}(H)}{\lambda_{\text{max}}(HKH)} \) results in

\[
\Delta V(\tilde{u}^{\text{ave}}) \leq -\frac{(\lambda_{\text{min}}(H))^2}{\lambda_{\text{max}}(HKH)} |\tilde{u}^{\text{ave}}|^2.
\]

(5.67)

For all \( \epsilon \in (0, 2\lambda_{\text{min}}(H)/\lambda_{\text{max}}(HKH)) \), the averaged system is exponentially stable, which completes the proof [5, Theorem 2.2.4]. \( \quad \Box \)

With the exponential stability of the averaged homogeneous error system established, we now consider the full system (5.34). First, from the perturbed averaged system

\[
\tilde{u}^{\text{ave}}(l + 1) = (I - \epsilon KH) \tilde{u}^{\text{ave}}(l) + \delta(l),
\]

(5.68)

where \( \delta(l) = [\delta_0(l), \ldots, \delta_{N-1}(l)]^T \), we see from the matrix \( K \) that \( \tilde{u}^{\text{ave}}(l) \) converges exponentially to an \( O(\max_k \{a_k\}) \)-neighborhood of the origin since from Lemma 5.2, \( |\delta_k(l)| \leq \epsilon^{-l} + \kappa_k \epsilon \tilde{u}^2 \). From [5], the exponential convergence rate of \( \tilde{u} \) in (5.34) tends to the rate of \( \tilde{u}^{\text{ave}} \) in the average system as \( \epsilon \) goes to zero. We can now state convergence results for the overall \( \tilde{u} \)-system and the cost value \( J \).
Theorem 5.2 Consider the full system (5.34) with the conditions of Theorem 5.1 satisfied. For sufficiently small $a_k$, $k \in \{0, \ldots, N - 1\}$, there exists $\epsilon_1^* \in (0, \epsilon^*]$, such that for all $\epsilon \in (0, \epsilon_1^*)$, the error variable $\tilde{u}$ locally exponentially converges to an $O(\max_k \{a_k\})$-neighborhood of the origin.

Corollary 5.1 With the conditions of Theorem 5.2 satisfied, the cost value $J$ locally exponentially converges to an $O(\max_i \{a_i^2\})$-neighborhood of the optimal cost $J(u^*)$.

Proof: Define $\tilde{J}(u) = J(u) - J(u^*)$. Then, from (5.23) and (5.25), we have

$$\tilde{J}(u) = \frac{1}{2}(u - u^*)^T H(u - u^*),$$

$$= \frac{1}{2}\tilde{u}^T H\tilde{u} + \mu^T H\tilde{u} + \frac{1}{2}\tilde{\mu}^T H\tilde{\mu}. \quad (5.69)$$

From Theorem 5.2, $\tilde{u}$ locally exponentially converges to an $O(\max_k \{a_k\})$-neighborhood of the origin. Thus, $\tilde{J}(u)$ locally exponentially converges to an $O(\max_i \{a_i^2\})$-neighborhood of the origin. ■

Finally, if the gains and phase values $a_i, \gamma_i, \varphi_i$ are chosen so that the elements of $K$ in (5.65) all equal $\nu > 0$, we can rewrite (5.65) as

$$\tilde{u}_k^{aw}(l + 1) = (I - \epsilon \nu H) \tilde{u}_k^{aw}(l) \quad (5.70)$$

and characterize the eigenvalues of $\Upsilon$.

Corollary 5.2 Consider the system (5.70) with $H > 0$ and scalars $\epsilon, \nu > 0$. The eigenvalues of $\Upsilon$ are given by

$$\lambda_i(\Upsilon) = 1 - \epsilon \nu \lambda_i(H), \quad i \in \{1, \ldots, N\}, \quad (5.71)$$

with bounds on $\lambda(H)$ given in Lemma 5.1.

Proof: Let $H = U T U^T$ be the Schur decomposition of $H$, where $T$ is upper triangular and $U$ is unitary. Then,

$$U^T \Upsilon U = U^T (I - \epsilon \nu H) U,$$

$$= I - \epsilon \nu T, \quad (5.72)$$

which is upper triangular with diagonal elements equal to (5.71). ■
5.5 Simulation Results

First, we present a scalar example to explore how the discrete-time system’s level of instability affects the optimization problem’s structure. As a barometer for the optimization problem’s difficulty, we compute the condition number of the Hessian, $\kappa(H)$. The remaining examples demonstrate convergence to the optimal control sequence and show how projection can be used to accommodate input constraints.

5.5.1 Example 1: Scalar System

Consider the scalar system

\[ x_{k+1} = \alpha x_k + u_k, \quad (5.73) \]

with initial condition $x_0 = 1$ and the cost function

\[ J = \frac{1}{2} x_N^2 + \frac{1}{2} \sum_{k=0}^{N-1} x_k^2 + u_k^2, \quad (5.74) \]

where $Q_N = Q_k = R_k = 1$. To study how the stability of (5.73) affects the eigenvalues of $H$, we will vary the parameter $\alpha$ in the interval $[0, 3]$ for intervals $N \in \{2, 3, 5, 8, 10\}$.

Figure 5.2 depicts how $\kappa(H)$ varies with $\alpha$ for each time horizon $N$, which indicates that the optimization problem becomes more difficult as the instability of the discrete-time system increases, and this difficulty is more acute for longer time horizons. In fact, $\kappa(H)$ appears to grow almost exponentially with $\alpha$ with a rate of increase dependent on the time horizon $N$. Figure 5.3 shows how, for this example, the eigenvalues appear to grow and spread apart until $\alpha > 1$ (at $\alpha = 1$, (5.73) is marginally stable) before converging to one—except for the maximal eigenvalue, which tracks $\kappa(H)$. Since $H$ is symmetric, $\kappa(H) = \lambda_{\text{max}}(H)/\lambda_{\text{min}}(H)$.

One should note that these structural changes to the Hessian are due to the inherent challenges of this optimization problem; they are not caused by the extremum seeking algorithm. Thus, if possible, one should consider attaining the optimal sequence $\{u_k^*\}_{k=0}^{N-1}$ recursively since $\kappa(H)$ grows with $N$, which means a
good initial estimate for the control sequence is more important for larger $N$. Specifically, first obtain $\{u_k^*\}_{k=0}^{M-1}$ and use $\{u_0^*, u_1^*, \ldots, u_{M-1}^*, u_M\}$ as the initial control sequence for $M + 1$ intervals, where two possible values for $u_M$ are $u_M = 0$ or $u_M = u_{M-1}^*$. Then, repeat this process until the desired $N$ intervals have been optimized. In practice, the current non-optimal control sequence may also be a good initial control sequence for the optimization scheme as well.

### 5.5.2 Example 2: Unstable Third-Order System

Consider the unstable system,

$$
x_{k+1} = \begin{bmatrix} 0.5 & 0.5 & 0.7 \\ 0.6 & 0.75 & -0.2 \\ 0.5 & 0 & 0.5 \end{bmatrix} x_k + \begin{bmatrix} -0.15 \\ 0.5 \\ 0.75 \end{bmatrix} u_k,
\quad (5.75)
$$

$$
y_k = \begin{bmatrix} 1 & 0.5 & 0 \\ 0.15 & -1.5 & 0 \end{bmatrix} x_k,
\quad (5.76)
$$

with initial condition $x_0 = [-1, -0.5, 0.1]^T$. The system’s eigenvalues are $-0.29$, $1.33$, and $0.71$. Also note that the third state does not directly appear in the system’s output.
Figure 5.3: Eigenvalues of $H$ versus $\alpha$ for $N = 5$ (top) and $N = 10$ (bottom). The maximum eigenvalue is approximately the condition number $\kappa(H) = \frac{\lambda_{\text{max}}(H)}{\lambda_{\text{min}}(H)}$, shown in Figure 5.2.

We want to minimize the cost function

$$ J = \frac{1}{2} \sigma_N y_N^2 + \frac{1}{2} \sum_{k=0}^{N-1} \sigma_k y_k^2 + u_k^2, $$

(5.77)

with penalty weights $\sigma_k$, when $u_k$ is unconstrained and when we have input constraints,

$$ u_k \leq u_k \leq \overline{u}_k. $$

(5.78)

For the unconstrained case, we implement the scheme (5.20)–(5.22), with $a_k = 0.1$, $\gamma_k = 0.1/(N - k)$, $\omega_k$ drawn from a uniform distribution such that $\omega_k \in (0, \pi)$, and $\varphi_k = -\angle W(e^{j\omega_k})$. The choice to scale $\gamma_k$ with $k$ is an attempt to mitigate the effect condition number $\kappa(H)$ as it increases with $N$. For the constrained case, we use the same parameters, but we modify the extremum seeking controller to accommodate the input constraints. Specifically, we implement

$$ \hat{u}_k(l + 1) = \hat{u}_k(l) - \text{Proj} \left\{ \gamma_k \xi_k(l); u_k, \overline{u}_k \right\}, $$

(5.79)

where the projection operator is defined as

$$ \text{Proj}\{\phi; u_k, \overline{u}_k\} = \begin{cases} 0, & \text{if } \hat{u}_k \leq u_k + \phi, \\ 0, & \text{if } \hat{u}_k \geq \overline{u}_k + \phi, \\ \phi, & \text{otherwise.} \end{cases} $$

(5.80)
Figure 5.4: Trajectory of $u_0$ and $u_1$ for Example 2 for the constrained (blue) and unconstrained (red) cases. The cost level sets are shown with the unconstrained minimum (green circle). The dashed lines denote the input constrained set with the optimum value on the boundary (cyan square).

This definition of the projection operator constrains $\hat{u}_k$ according to the input constraints. The input value $u_k$, however, is allowed to violate the constraints, but only by the perturbation magnitude $a_k$. In this example, we study the input bounds $|u_k| \leq 1$.

For both the unconstrained and constrained scenarios, we simulate the system with a zero initial control sequence for the finite horizon $N = 3$ and penalty weights $\sigma_k = 1$ for all $k$. Figure 5.4 depicts the planar trajectories of $u_0$ and $u_1$ for both cases with the level sets of $J(u)$ and the input constraints (dashed lines) superimposed. In the example, the constrained trajectory hits the upper limit of $\hat{u}_1$ and climbs along the boundary until reaching the constrained optimum. The cyan square denotes the optimal control values for the constrained problem and is obtained by minimizing $J(u)$, subject to the input constraints, by formulating a quadratic program [26]. Figure 5.5 depicts the time history of $u_2$ in both cases. The optimal value is denoted by a black line. For completeness, $\kappa(H) = 2.75$ when $N = 3$. 
Figure 5.5: Time history of $u_2$ for both the constrained (top) and unconstrained (bottom) cases in Example 2. Black lines denote the optimal values.

5.6 Summary

We have introduced a non-model based approach to solve the finite-time horizon optimal control problem for unknown discrete-time systems. This method attains the open-loop control sequence $\{u^*_k\}_{k=0}^{N-1}$ that minimizes a cost function quadratic in the input and output and extends the convergence results found in [23]. Such a framework may be a natural tool for the real-time optimization of highly repetitive systems that are unknown and potentially high-dimensional.

The convergence result does not depend on the time horizon length $N$ or on the stability of the unknown system (because the time horizon is finite). The optimization problem, independent of the solver, becomes more difficult as $N$ increases and as the instability of the system increases. By using a Newton-based extremum seeking controller [42], rather than the gradient descent-based, a convergence rate independent of the Hessian is possible and may facilitate the optimization of problems with longer time horizons. However, an initial estimate of the Hessian is required.


The dissertation author is the primary investigator and author of this work.
Chapter 6

Leader-Enabled Deployment to Planar Curves

We introduce a framework for multi-agent deployment into families of geometric curves. Our design employs linear reaction-advection-diffusion equations and boundary control techniques, which treat the agents as a continuum. These PDE models are an application of the internal model principle [35], but in a spatial sense, where the PDE models allow the agents to achieve a family of deployments that correspond to the models’ nonzero equilibria. Specifically, the follower agents’ feedback laws, incorporated by the PDE models, contain a model for a family of geometric curves, namely, signals in space rather than time. Consequently, the agents are able to achieve different formations while using the same controller, whereas other works require a controller’s parameters to be changed to achieve a new deployment.

The follower agents’ feedback laws, however, do not ensure stability, and many, if not most, non-elementary deployments are derived from unstable PDE models. Other PDE-based designs utilize only inherently stable models. Stabilization of the planar curves is guaranteed by two boundary agents—the leader agent and the anchor agent—which serve as boundary conditions for the PDE models. These special agents execute control laws designed using the backstepping approach [96]. Even for standard deployments, which correspond to stable PDEs, such as rendezvous or deployment to a line, our leader feedback can achieve any
desired convergence speed, in contrast to the convergence speeds of the standard consensus-based algorithms that are limited by the first eigenvalue of the heat equation [19, 32, 33]. The desired deployment shape is encoded in these boundary controls in the form of a bias term, which allows the leader and anchor to select a specific curve from the deployment family for the agents to stabilize. By adjusting their respective bias terms, the leader and anchor (and by extension, the user) can induce the agents to unknowingly deploy to other curves within the deployment family.

Our framework also includes the design of observers, which are employed by the leader agent, using the backstepping approach for PDEs with boundary sensing [97]. These observers require knowledge of only the position of the leader’s nearest neighbor to estimate the positions of all the agents. By spatially discretizing the PDE models, observers, and boundary controllers, we obtain control laws for the follower agents, the anchor, and the leader. This discretization imposes a fixed communication topology on the agents, shown in Figure 6.1, where if the leader employs output feedback, all the agents utilize only local information (in the sense of the communication topology) in their control laws.

This chapter is organized as follows: Section 6.1 introduces leader-enabled deployment for both decoupled 1-D deployments and complex-valued deployments with $x$-$y$ cross-coupling, Section 6.2 details the design of stabilizing controllers with closed-loop stability proven in Section 6.3, Section 6.4 analyzes the observer design, Section 6.5 discretizes the infinite-dimensional controllers for $N$ agents, Section 6.6 presents numerical simulations, and Section 6.7 provides a summary of the results.

### 6.1 Leader-Enabled Deployment

For the planar deployment problem, we consider a large (continuum) group of fully actuated agents operating in a common reference frame, namely, we consider the dynamical model,

$$x_t(\alpha, t) = u(\alpha, t),$$  \hspace{1cm} (6.1)

$$y_t(\alpha, t) = v(\alpha, t),$$  \hspace{1cm} (6.2)
Figure 6.1: Communication topology imposed by spatial discretization. The leader needs global information (dashed) unless it uses output feedback. Legend: triangle = anchor, plus = follower, star = leader.

where \((x(\alpha, t), y(\alpha, t))\) denotes the position of agent \(\alpha\) at time \(t\), \(u(\alpha, t)\) and \(v(\alpha, t)\) are the control inputs for agent \(\alpha\), \(x_t(\alpha, t) = \frac{\partial}{\partial t} x(\alpha, t)\), \(y_t(\alpha, t) = \frac{\partial}{\partial t} y(\alpha, t)\), and \(\alpha \in [0, 1]\). We refer to the parameter \(\alpha\) as the agent identity, which serves as an agent’s identification number and as the spatial variable of a PDE model for the group’s collective dynamics. Discretizing (6.1)–(6.2) with respect to \(\alpha\), leads to the following dynamical model for agent \(i\):

\[
\begin{align*}
\dot{x}_i(t) &= u_i(t), \\
\dot{y}_i(t) &= v_i(t),
\end{align*}
\]

where \(i \in \{0, \ldots, n\}\). For later use, we define the notation: \(x_\alpha(\alpha, t) = \frac{\partial}{\partial \alpha} x(\alpha, t)\), \(x_{\alpha \alpha}(\alpha, t) = \frac{\partial^2}{\partial \alpha^2} x(\alpha, t)\).

Our goal is to stably deploy a continuum of agents to families of planar curves by designing the controllers \(u(\alpha, t), v(\alpha, t)\). Then, a finite number of agents implement the discretized controllers \(u_i(t), v_i(t)\). This 2-D deployment problem can be approached either through (A) two decoupled 1-D deployment problems, where the horizontal feedback is decoupled from vertical feedback, i.e., actuation in the \(x\)-direction does not depend on the position measurement in the \(y\)-direction, and vice-versa, or (B) as a single complex-valued deployment, where the real and imaginary components represent the horizontal and vertical coordinates and actuation in each
coordinate direction depends on the entire position vector. Restated, in (A), the horizontal velocity command is a function of only the \(x\)-position and the vertical velocity command is a function of only the \(y\)-position. In (B), the complex-valued formulation allows for horizontal and vertical velocity commands that are functions of the planar position \((x,y)\). Using two decoupled 1-D deployments is simpler than employing a single complex-valued deployment, so we consider it first for clarity.

6.1.1 Decoupled 1-D Deployments

It is common to approach the deployment problem through consensus-based control laws [19, 32, 60, 65], whose basic form is given by

\[
\dot{x}_i(t) = \sum_{j \in N_i} (x_j(t) - x_i(t)),
\]

(6.5)

where \(N_i\) denotes the set of agents/neighbors that communicate with agent \(i\). In [33], (6.5) is formally shown to coincide with the heat equation,

\[
x_t(\alpha, t) = x_{\alpha\alpha}(\alpha, t),
\]

(6.6)

where each agent employs the diffusion-based feedback \(u(\alpha, t) = x_{\alpha\alpha}(\alpha, t)\), which depends only on local agent interactions, i.e., an agent’s nearest neighbors. This simple agent strategy is stable, but it is limited in its convergence rate and is capable of achieving only linear formations (because the equilibrium equation is the simplest second order ODE, \(\ddot{x}(\alpha) = 0\)).

Remark 6.1 Throughout this work, “nearest-neighbor” refers to agents that are nearest in terms of the fixed communication topology, not in terms of physical distance.

Drawing from the connection between consensus and the heat equation, we approach the PDE-based deployment with the more general linear reaction-advection-diffusion equation,

\[
x_t(\alpha, t) = x_{\alpha\alpha}(\alpha, t) + bx_\alpha(\alpha, t) + \lambda x(\alpha, t),
\]

(6.7)
Table 6.1: Basis functions for 1-D deployment curves of the reaction-advection-diffusion equation.

<table>
<thead>
<tr>
<th>$b, \lambda$ Values</th>
<th>Basis Functions $(\psi_0(\alpha), \psi_1(\alpha))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = \lambda = 0$</td>
<td>$(1, \alpha)$</td>
</tr>
<tr>
<td>$b \neq 0, \lambda = 0$</td>
<td>$(1, e^{-b\alpha})$</td>
</tr>
<tr>
<td>$b = 0, \lambda &gt; 0$</td>
<td>$(\cos(\theta\alpha), \sin(\theta\alpha)), \theta = \sqrt{\lambda}$</td>
</tr>
<tr>
<td>$b = 0, \lambda &lt; 0$</td>
<td>$(\cosh(\sigma\alpha), \sinh(\sigma\alpha)), \sigma = \sqrt{-\lambda}$</td>
</tr>
<tr>
<td>$b^2 = 4\lambda$</td>
<td>$(e^{-\sigma\alpha}, \alpha e^{-\sigma\alpha}), \sigma = \sqrt{\lambda}$</td>
</tr>
<tr>
<td>$b^2 &lt; 4\lambda$</td>
<td>$(e^{\sigma\alpha} \cos(\theta\alpha), e^{\sigma\alpha} \sin(\theta\alpha)), \sigma = -\frac{b}{2}, \theta = \frac{1}{2} \sqrt{4\lambda - b^2}$</td>
</tr>
<tr>
<td>$b^2 &gt; 4\lambda$</td>
<td>$(e^{\sigma_0\alpha}, e^{\sigma_1\alpha}), \sigma_0 = -\frac{b}{2} + \frac{1}{2}\sqrt{b^2 - 4\lambda}, \sigma_1 = -\frac{b}{2} - \frac{1}{2}\sqrt{b^2 - 4\lambda}$</td>
</tr>
</tbody>
</table>

where the agents’ velocity-actuated feedback laws $u(\alpha, t)$ are given by the right-hand side of (6.7), and $v(\alpha, t)$ follows analogously for the $y$-dimension. These feedback laws maintain the simplicity of the diffusion-based feedback as they are still based on nearest-neighbor information with all the agents applying the same constant gains $b$ and $\lambda$. In the sequel, we drop the arguments $(\alpha, t)$ whenever the context allows us to do so without harming clarity.

We designate a special role for the two boundary agents, i.e., agent $\alpha = 0$ and agent $\alpha = 1$, whose motions are governed by

$$x_t(0, t) = U_0(t), \quad (6.8)$$
$$x_t(1, t) = U_1(t), \quad (6.9)$$

where $U_0(t)$ and $U_1(t)$ are controls to be designed, and which act as the boundary conditions for the PDE (6.7). The **leader** agent ($\alpha = 1$) and the **anchor** agent ($\alpha = 0$) will control the follower agents ($0 < \alpha < 1$). As indicated by their names, the leader stabilizes the deployment profile $\bar{x}(\alpha)$ while the anchor simply deploys to its designated position $\bar{x}(0)$. Either the leader agent or both the leader and anchor agents may be treated as virtual agents if desired (and as suggested by the use of virtual edge leaders in [65]), but it is not necessary.
The deployment families of interest correspond to the nonzero equilibrium curves of (6.7), which satisfy the two-point boundary value problem,

$$\ddot{x}(\alpha) + b\dot{x}(\alpha) + \lambda x(\alpha) = 0,$$

(6.10)

with $x(0)$ and $x(1)$ given. This allows for a much more general family of deployments than the linear (in $\alpha$) equilibrium curves of the heat equation (6.6). Equation (6.10), which is a second-order ODE with constant coefficients, characterizes all the achievable 1-D deployments with the follower agent feedbacks (6.7). While these feedbacks make these deployments feasible, they do not guarantee stability since the open-loop response of (6.7) is $x(\alpha, t) = \sum_{k=1}^{\infty} C_k e^{\chi t} \sin(k \pi \alpha)$ with eigenvalues $\chi = \lambda - b^2/4 - \pi^2 k^2$, $k \in \{1, 2, \ldots\}$, and constants $C_k$, whose values depend on the initial condition $x(\alpha, 0)$. In particular, deployment families where $\lambda > b^2/4 + \pi^2$ are unstable. Hence, the leader and the anchor agents play a crucial role in stabilizing the possibly nonlinear (in $\alpha$) deployment curves.

For planar deployment, we utilize a 1-D PDE model for each coordinate axis, which yields two deployments, $\bar{x}(\alpha)$ and $\bar{y}(\alpha)$, that characterize a planar curve parameterized in $\alpha$,

$$\bar{x}(\alpha) = a_0 \psi_0(\alpha) + a_1 \psi_1(\alpha), \quad \bar{y}(\alpha) = a_2 \psi_2(\alpha) + a_3 \psi_3(\alpha),$$

(6.11)

where $(\psi_0, \psi_1)$ and $(\psi_2, \psi_3)$ are basis functions associated with the solutions of (6.10) for the respective horizontal and vertical PDE models. We term the coefficients, $a_0$, $a_1$, $a_2$, and $a_3$, the deployment coefficients, which are scalars the user is free to select to define a desired deployment. It is of interest to see how rich the family of possible geometric curves is. Table 6.1 categorizes the basis functions according to the values of $b$ and $\lambda$. To the user, who has particular planar formations in mind, the basis functions are a starting point in selecting the strategies of the follower agents, and also of the leader and anchor agents.

The ability to use two disparate PDE models (one for each dimension) provides the user with a wide variety of basis function combinations that produce various planar deployments. Interestingly, the well-known Lissajous curves, given
Figure 6.2: Lissajous curves for various values of $C$, $D$, and $E$ with $A = B = 1$, $\alpha \in [0, 1]$. The dots represent agents. Legend: black = 1 agent, red = 2 agents.

by

\[ \bar{x}(\alpha) = A \sin(C\alpha + E), \]
\[ \bar{y}(\alpha) = B \sin(D\alpha), \]

where $A$, $B$, $C$, $D$, and $E$ are scalars with $C, D > 0$, are achieved when the 1-D deployments are governed by the reaction-diffusion equations, $x_t = x_{\alpha\alpha} + \sqrt{C}x$ and $y_t = y_{\alpha\alpha} + \sqrt{D}y$, and the following deployment coefficients are selected: $a_0 = A \sin E$, $a_1 = A \cos E$, $a_2 = 0$, and $a_3 = B$. Figure 6.2 depicts four possible deployments of 15 agents based on Lissajous curves.

When the same PDE model is used in each dimension, the parameterized deployment can be written in vector form as,

\[
\begin{bmatrix}
\bar{x}(\alpha) \\
\bar{y}(\alpha)
\end{bmatrix}
= \begin{bmatrix}
a_0 & a_1 \\
a_2 & a_3
\end{bmatrix}
\begin{bmatrix}
\psi_0(\alpha) \\
\psi_1(\alpha)
\end{bmatrix},
\]

(6.12)

and the coefficient matrix can be chosen to be a rotation, scaling, shear, or reflection matrix. For example, the coefficients can be selected to define the desired deployment,

\[
\begin{bmatrix}
\bar{x}(\alpha) \\
\bar{y}(\alpha)
\end{bmatrix}
= \begin{bmatrix}
\cos \phi & \sin \phi \\
-\sin \phi & \cos \phi
\end{bmatrix}
\begin{bmatrix}
b_0 \psi_0(\alpha) \\
b_1 \psi_1(\alpha)
\end{bmatrix},
\]

(6.13)

which is a counterclockwise rotation of the scaled curve $\bar{x}(\alpha) = b_0 \psi_0(\alpha)$, $\bar{y}(\alpha) = b_1 \psi_1(\alpha)$ about the origin by the angle $\phi$. When two identical reaction-diffusion
equations with $\lambda > 0$ are used, $(\psi_0(\alpha), \psi_1(\alpha)) = (\cos(\sqrt{\lambda} \alpha), \sin(\sqrt{\lambda} \alpha))$ and (6.13) represents a rotated ellipse. If the same PDE models are used but with $\lambda < 0$, $(\psi_0(\alpha), \psi_1(\alpha)) = (\cosh(\sqrt{-\lambda} \alpha), \sinh(\sqrt{-\lambda} \alpha))$ and (6.13) represents a rotated hyperbola.

6.1.2 Complex-valued 2-D Deployment

Until now, we have considered the already rich family of planar deployments that are created by pairing two independent 1-D deployments. We now extend this family of achievable deployments by utilizing an agent’s full position vector $z = (x, y)$ in a feedback law for each coordinate direction. To do so, we consider the complex-valued Ginzburg-Landau PDE as a continuum model of the collective dynamics of the agents in the plane.

Let $z(\alpha, t) = x(\alpha, t) + jy(\alpha, t)$ be the complex-valued position at time $t$ of agent $\alpha$ where $j$ denotes the imaginary unit, $\sqrt{-1}$. Now consider the complex-valued reaction-advection-diffusion equation (which is a linear Ginzburg-Landau PDE with constant coefficients),

$$z_t(\alpha, t) = (\varepsilon R + j \varepsilon I)z_{\alpha\alpha}(\alpha, t) + (b R + j b I)z_\alpha(\alpha, t) + (\lambda R + j \lambda I)z(\alpha, t),$$

(6.14)

where $\varepsilon_R > 0$. In the sequel, we use $\varepsilon = \varepsilon_R + j \varepsilon_I$, $b = b_R + j b_I$, and $\lambda = \lambda_R + j \lambda_I$ for conciseness. Equation (6.14) represents the followers’ velocity-actuated feedback laws. As before, the leader and anchor agents serve as the boundary conditions for the PDE (6.14),

$$z_t(0, t) = U_0(t),$$

(6.15)

$$z_t(1, t) = U_1(t),$$

(6.16)

where $U_0(t)$ and $U_1(t)$ are controls to be designed. The open-loop system is unstable when $\lambda_R$ is positive and large.

The deployments associated with (6.14) are the equilibrium curves that satisfy the complex-valued two-point boundary value problem,

$$\varepsilon \bar{z}''(\alpha) + b \bar{z}'(\alpha) + \lambda \bar{z}(\alpha) = 0,$$

(6.17)
where $\bar{z}(0)$ and $\bar{z}(1)$ are given. The second-order complex ODE (6.17) is in fact a fourth order real ODE, whose solution is given in terms of four basis functions as $\bar{x}(\alpha) + j\bar{y}(\alpha) = (a_0 + ja_1) (\psi_0(\alpha) + j\psi_1(\alpha)) + (a_2 + ja_3) (\psi_2(\alpha) + j\psi_3(\alpha))$, alternatively written as

$$\begin{bmatrix} \bar{x}(\alpha) \\ \bar{y}(\alpha) \end{bmatrix} = \begin{bmatrix} a_0 & -a_1 \\ a_1 & a_0 \end{bmatrix} \begin{bmatrix} \psi_0(\alpha) \\ \psi_1(\alpha) \end{bmatrix} + \begin{bmatrix} a_2 & -a_3 \\ a_3 & a_2 \end{bmatrix} \begin{bmatrix} \psi_2(\alpha) \\ \psi_3(\alpha) \end{bmatrix}. \quad (6.18)$$

The presence of four basis functions affords the designer additional flexibility in shaping deployments, but the restrictive structure of the matrices in (6.18) prevents the user from being able to shear, reflect, or scale disproportionately the formations. The deployments can only be rotated or equally scaled.

For the second-order complex-valued ODE (6.17), the resulting basis functions are not easy to categorize in terms of the values of the real and imaginary parts of $\varepsilon$, $b$, $\lambda$, as was done in Table 6.1 for real second-order ODEs. We characterize the basis functions for specific subclasses of the complex-valued reaction-advection-diffusion equation. For $b = \lambda = 0$, the equilibrium profiles are linear in $\alpha$, regardless of the value of $\varepsilon$. The more interesting cases are characterized next.

**Advection-diffusion equation ($\lambda = 0$)**

$$\begin{bmatrix} \bar{x}(\alpha) \\ \bar{y}(\alpha) \end{bmatrix} = \begin{bmatrix} a_0 & -a_1 \\ a_1 & a_0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} a_2 & -a_3 \\ a_3 & a_2 \end{bmatrix} \begin{bmatrix} \sin(\theta \alpha) \\ \cos(\theta \alpha) \end{bmatrix} e^{-\sigma \alpha}, \quad (6.19)$$

where $\theta = (\varepsilon_R b_I - b_R \varepsilon_I)/|\varepsilon|^2$, $\sigma = (\varepsilon_R b_I + \varepsilon_I b_R)/|\varepsilon|^2$, and $|\varepsilon| = \sqrt{\varepsilon_R^2 + \varepsilon_I^2}$. If $\varepsilon_R b_I = b_R \varepsilon_I$, the deployment reduces to two, independent 1-D profiles.

**Reaction-diffusion equation ($b = 0$)**

$$\begin{bmatrix} \bar{x}(\alpha) \\ \bar{y}(\alpha) \end{bmatrix} = \begin{bmatrix} a_0 & -a_1 \\ a_1 & a_0 \end{bmatrix} \begin{bmatrix} \sin(\theta \alpha) \sinh(\sigma \alpha) \\ \cos(\theta \alpha) \cosh(\sigma \alpha) \end{bmatrix} + \begin{bmatrix} a_2 & -a_3 \\ a_3 & a_2 \end{bmatrix} \begin{bmatrix} \sin(\theta \alpha) \cosh(\sigma \alpha) \\ \cos(\theta \alpha) \sinh(\sigma \alpha) \end{bmatrix}, \quad (6.20)$$
where
\[
\theta = \frac{1}{\sqrt{2}|\varepsilon|} \sqrt{|\varepsilon||\lambda| + \varepsilon_R\lambda_R + \lambda_I\varepsilon_I},
\]
\[
\sigma = \frac{1}{\sqrt{2}|\varepsilon|} \text{sgn}(\varepsilon_R\lambda_I - \varepsilon_I\lambda_R) \sqrt{|\varepsilon||\lambda| - \varepsilon_R\lambda_R - \lambda_I\varepsilon_I},
\]
and sgn denotes the signum function. The deployments become decoupled if \(\varepsilon_R\lambda_I = \lambda_R\varepsilon_I\).

Reaction-advection-diffusion equation
\[
\begin{bmatrix}
\bar{x}(\alpha) \\
\bar{y}(\alpha)
\end{bmatrix} =
\begin{bmatrix}
a_0 & -a_1 \\
a_1 & a_0
\end{bmatrix}
\begin{bmatrix}
\cos(\theta_0\alpha) \\
\sin(\theta_0\alpha)
\end{bmatrix} e^{\sigma_0\alpha}
+ \begin{bmatrix}
a_2 & -a_3 \\
a_3 & a_2
\end{bmatrix}
\begin{bmatrix}
\cos(\theta_1\alpha) \\
\sin(\theta_1\alpha)
\end{bmatrix} e^{\sigma_1\alpha},
\] (6.21)

where
\[
\theta_0 = -\frac{1}{2\sqrt{2}|\varepsilon|^2} \left( \sqrt{2}(\varepsilon_R b_I - \varepsilon_I b_R) - \varepsilon_R\text{sgn}(a_I)\sqrt{|a| - a_R + \varepsilon_I\sqrt{|a| + a_R}} \right),
\]
\[
\theta_1 = -\frac{1}{2\sqrt{2}|\varepsilon|^2} \left( \sqrt{2}(\varepsilon_R b_I - \varepsilon_I b_R) + \varepsilon_R\text{sgn}(a_I)\sqrt{|a| - a_R - \varepsilon_I\sqrt{|a| + a_R}} \right),
\]
\[
\sigma_0 = -\frac{1}{2\sqrt{2}|\varepsilon|^2} \left( \sqrt{2}(e_R b_R + e_I b_I) - \varepsilon_R\sqrt{|a| + a_R - \varepsilon_I\text{sgn}(a_I)\sqrt{|a| - a_R}} \right),
\]
\[
\sigma_1 = -\frac{1}{2\sqrt{2}|\varepsilon|^2} \left( \sqrt{2}(e_R b_R + e_I b_I) + \varepsilon_R\sqrt{|a| + a_R + \varepsilon_I\text{sgn}(a_I)\sqrt{|a| - a_R}} \right),
\]
\[
a_R = b_R^2 - b_I^2 - 4(\varepsilon_R\lambda_R - \varepsilon_I\lambda_I),
\]
\[
a_I = 2b_R b_I - 4(\varepsilon_R\lambda_I + \varepsilon_I\lambda_R),
\]
and \(|a| = \sqrt{a_R^2 + a_I^2} \).

The deployments in each dimension become decoupled if \(|a| = 0\).

Clearly, selecting the appropriate PDE model (6.14), specifically the coefficients \(\varepsilon, b, \) and \(\lambda\) required for a desired deployment family, is not as straightforward as in the 1-D case due to the complicated expressions for \(\theta\) and \(\sigma\) given in (6.19), (6.20), and (6.21). However, interesting deployments can be found with some effort. For example, by selecting \(\varepsilon = 1/2 + j/2, b = \pi - j\pi, \) and \(\lambda = 0\), the deployment (6.19) becomes
\[
\begin{bmatrix}
\bar{x}(\alpha) \\
\bar{y}(\alpha)
\end{bmatrix} =
\begin{bmatrix}
a_0 & -a_1 \\
a_1 & a_0
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix} + \begin{bmatrix}
a_2 & -a_3 \\
a_3 & a_2
\end{bmatrix}
\begin{bmatrix}
-sin(2\pi\alpha) \\
\cos(2\pi\alpha)
\end{bmatrix},
\] (6.22)
which represents a circle deployment centered about the point \((a_0, a_1)\). By simply changing the scalars \(a_0\) and \(a_1\), the deployment can be moved about the plane. In contrast, a circle deployment can also be formed using two, independent, open-loop unstable 1-D reaction-diffusion equations with \(\lambda = 4\pi^2\), whose equilibria correspond to a family of ellipses centered about the origin. The deployments (6.20) and (6.21) model families of spiral-like deployments.

### 6.1.3 Design Procedure for Desired Deployment Profiles

To achieve leader-enabled deployment onto possibly nonlinear (in \(\alpha\)) planar curves using either two 1-D deployments or one complex-valued deployment, the user applies the following steps:

1. Select a desired deployment family, i.e., a family of basis functions.
2. Select specific basis functions by choosing the coefficients of the appropriate PDE model(s).
3. Choose deployment coefficients to generate a specific planar deployment.
4. Choose the desired deployment convergence rate.
5. Discretize the PDE model(s) spatially to obtain implementable control laws for the leader, anchor, and follower agents.

This procedure encompasses both feasibility (steps 1–3) and stability (steps 4–5). We have discussed deployment families for agents with feedback laws derived from PDE model(s), which guarantee feasibility, but not stability, of the deployments. For stable deployment, we design the control laws \(U_0(t)\) and \(U_1(t)\) for the anchor and the leader, and for the leader, observers to estimate the agents’ positions. We now focus on these designs.

### 6.2 Leader Feedback Design

We employ PDE backstepping boundary control [96] for the complex-valued PDE model (6.14)–(6.16) since leader-based control naturally leads to formulations
with actuation at the boundary. PDE backstepping succeeds in deriving closed-form controllers that achieve exponential stability with only boundary actuation. This approach is more elegant than other boundary methods that produce complicated controllers, which require solving operator Riccati equations. This design also applies, as a special case, to the real-valued PDE model (6.7) with boundary conditions (6.8), (6.9) by treating \( z(\alpha, t), \varepsilon, b, \) and \( \lambda \) as real-valued and setting \( \varepsilon = 1. \)

First, we introduce the deployment profile error,

\[
\zeta(\alpha, t) = [x(\alpha, t) + jy(\alpha, t)] - [\bar{x}(\alpha) + j\bar{y}(\alpha, t)],
\]

\[
= z(\alpha, t) - \bar{z}(\alpha, t),
\]

(6.23)

to shift the equilibrium of (6.14) to the origin. Substituting (6.23) into (6.14)–(6.16) yields

\[
\zeta_t = \varepsilon\zeta_{\alpha\alpha} + b\zeta_{\alpha} + \lambda\zeta,
\]

(6.24)

\[
\zeta_t(0) = U_0(t),
\]

(6.25)

\[
\zeta_t(1) = U_1(t),
\]

(6.26)

and we remind the reader that the coefficients \( \varepsilon, b, \) and \( \lambda \) are complex and \( \text{Re}\{\varepsilon\} > 0. \) Next, we substitute the change of variable [96]

\[
v(\alpha) = \zeta(\alpha)e^{\frac{b}{4\varepsilon}\alpha},
\]

(6.27)

into (6.24)–(6.26) to eliminate the advection term and obtain

\[
v_t = \varepsilon v_{\alpha\alpha} + \left( \lambda - \frac{b^2}{4\varepsilon} \right) v,
\]

(6.28)

\[
v_t(0) = U_0(t),
\]

(6.29)

\[
v_t(1) = e^{\frac{b}{4\varepsilon}}U_1(t).
\]

(6.30)

Now let \( w(\alpha, t) \) be a new state that is defined by the coordinate transformation

\[
w(\alpha, t) = v(\alpha, t) - \int_{0}^{\alpha} k(\alpha, \beta)v(\beta, t) \, d\beta,
\]

(6.31)
where the kernel $k(\alpha, \beta)$ defined on $T = \{(\alpha, \beta) : 0 \leq \beta \leq \alpha \leq 1\}$ is given by [96]

$$k(\alpha, \beta) = -\gamma \beta \frac{I_1\left(\sqrt{\gamma \left(\alpha^2 - \beta^2\right)}\right)}{\sqrt{\gamma \left(\alpha^2 - \beta^2\right)}}, \quad (6.32)$$

with $\gamma = \left(c + \lambda - \frac{b^2}{4\varepsilon}\right)/\varepsilon$, $c > 0$. In (6.32), $I_1$ denotes the first order modified Bessel function of the first kind. The variable (6.31) is shown to satisfy the target PDE system,

$$w_t = \varepsilon w_{\alpha\alpha} - cw - \varepsilon k_\beta(\alpha,0)w(0), \quad (6.33)$$
$$w_t(0) = -cw(0), \quad (6.34)$$
$$w_t(1) = -cw(1), \quad (6.35)$$

and this transformation can be inverted to obtain

$$v(\alpha, t) = w(\alpha, t) + \int_0^\alpha l(\alpha, \beta) w(\beta, t) d\beta, \quad (6.36)$$

where the inverse gain kernel $l \in C^2(T)$ [96]. As will be seen in Section 6.3, the parameter $c$, which is selected by the user, determines the convergence rate of the deployment.

From (6.23), (6.27), (6.31), and the boundary conditions (6.29), (6.34), we obtain the anchor’s control law,

$$U_0(t) = -c\zeta(0, t) = -cz(0, t) + c\bar{z}(0). \quad (6.37)$$

For the leader’s control law, we introduce the operator $\mathcal{K}\{\cdot\}$ acting on the function $\xi(\alpha, t)$ as

$$\mathcal{K}\{\xi\}(t) = \varepsilon\sqrt{\gamma} I_1(\sqrt{\gamma}) e^{-b\sqrt{\gamma}} \xi(0, t) + \left(c + \frac{\varepsilon\gamma^2}{8} - \frac{\varepsilon\gamma}{2} + \frac{b\gamma}{4}\right) \xi(1, t)$$
$$+ \varepsilon\frac{\gamma}{2} \xi(1, t) + \varepsilon\gamma^2 \int_0^1 e^{-\frac{b}{\varepsilon\gamma}(1-\beta)} \Pi(\beta) \xi(\beta, t) d\beta, \quad (6.38)$$

$$\Pi(\beta) = \frac{\gamma^3}{(\gamma(1-\beta^2))^{\frac{3}{2}}} \left[3\beta \left(\sqrt{\frac{\gamma(1-\beta^2)}{(\gamma(1-\beta^2))^{\frac{3}{2}}}}\right) + \beta \left(\frac{\gamma^2}{\gamma(1-\beta^2)}\right)\right], \quad (6.39)$$
where \( I_2 \) and \( I_3 \) indicate the second and third order modified Bessel functions of the first kind respectively. From (6.23), (6.27), (6.31), and the boundary conditions (6.30), (6.35), we arrive at the leader’s control law,

\[
U_1(t) = -\mathcal{K}\{\zeta\}(t) = -\mathcal{K}\{z\}(t) + \mathcal{K}\{\bar{z}\},
\]

that, with (6.37), stabilizes the deployment profile \( \bar{z}(\alpha) \).

Of note, the control laws (6.37) and (6.40) both contain a feedback term and a constant bias term, whose value is determined by the desired formation and can be computed prior to deployment. By simply changing the bias terms—without changing the feedback terms or the follower agents’ control strategy—different deployment profiles can be induced and stabilized by the anchor and the leader. To achieve a specific formation, the user selects the deployment coefficients \( a_0, \ldots, a_3 \) to compute the bias \( \mathcal{K}\{\bar{z}\} \). If the bias terms are zero, rendezvous at the origin is achieved.

However, if the user has no knowledge of the deployment family and instead, changes the bias values directly, i.e., employs the boundary conditions

\[
z_t(0) = -cz(0) + \bar{U}_0, \\
z_t(1) = -\mathcal{K}\{z\}(t) + \bar{U}_1,
\]

where \( \bar{U}_0, \bar{U}_1 \in \mathbb{C} \) denote bias terms set by the user, the agents will stabilize the deployment profile,

\[
\bar{z}(\alpha) = e^{-\frac{b}{\bar{\varepsilon}}\alpha} \left( \bar{w}(\alpha) + \int_0^\alpha l(\alpha, \beta) \bar{w}(\beta) d\beta \right).
\]

This profile is found by applying the change of variable, \( v(\alpha, t) = z(\alpha, t)e^{\frac{b}{\bar{\varepsilon}}\alpha} \), and the transformation (6.31) to the system (6.14)–(6.16) to obtain the target PDE system (6.33) with boundary conditions, \( w_t(0, t) = -cw(0, t) + \bar{U}_0 \) and \( w_t(1, t) = -cw(1, t) + e^{\frac{b}{\bar{\varepsilon}}} \bar{U}_1 \). The equilibrium of (6.33) with these boundary conditions is

\[
\bar{w}(\alpha) = \frac{e^{\frac{b}{\bar{\varepsilon}}\bar{U}_1}}{c\sinh \Theta} \sinh (\Theta\alpha) + \frac{\bar{U}_0}{c} \left( \cosh (\Theta\alpha) - \frac{\cosh \Theta}{\sinh \Theta} \sinh (\Theta\alpha) \right) \\
- \frac{\sqrt{\gamma}\bar{U}_0}{c\Theta} \int_0^\alpha \frac{\sinh (\Theta(\alpha - \tau))}{\tau} I_1 (\sqrt{\gamma}\tau) \, d\tau \\
+ \frac{\sqrt{\gamma}\bar{U}_0 \sinh (\Theta\alpha)}{c\Theta \sinh \Theta} \int_0^1 \frac{\sinh (\Theta(1 - \tau))}{\tau} I_1 (\sqrt{\gamma}\tau) \, d\tau,
\]

(6.41)
where $\Theta = \sqrt{\frac{\varepsilon R}{\varepsilon I} j \varepsilon i} / |\varepsilon|$. We use the inverse transformation (6.36) and change of variable, $v(\alpha, t) = z(\alpha, t) e^{\frac{b}{2\varepsilon^2} \alpha}$, to obtain $\bar{z}(\alpha)$.

### 6.3 Closed-Loop Stability

Due to the dynamic character of the boundary conditions (6.34), (6.35), there are several aspects in which the stability analysis here differs from [96] and other work on PDE boundary control. A boundary value-dependent perturbation term arises on the right-hand side of (6.33) and the dynamic boundary conditions necessitate that the analysis be conducted in the Sobolev space,

$$H^1 = \left\{ f \in C^1[0, 1] \mid \|f\|^2 = |f(0)|^2 + |f(1)|^2 + \int_0^1 |f(\alpha)|^2 d\alpha + \int_0^1 |f_\alpha(\alpha)|^2 d\alpha < \infty \right\},$$

rather than the function space,

$$L^2 = \left\{ f \in C[0, 1] \mid \|f\|^2 = \int_0^1 |f(\alpha)|^2 d\alpha < \infty \right\}.$$ (6.43)

**Theorem 6.1** The system (6.14) with boundary conditions (6.15), (6.16) and control laws (6.37), (6.40) is exponentially stable in the $H^1$ norm at the equilibrium $z(\alpha, t) \equiv \bar{z}(\alpha)$, i.e., there exists $M > 0$ such that for all $t \geq 0$,

$$\Omega(t) \leq Me^{-ct}\Omega(0),$$

where $c > 0$, $c \neq b^2/(4\varepsilon^2) - \lambda$, and

$$\Omega(t) = |\zeta(0, t)|^2 + |\zeta(1, t)|^2 + \int_0^1 |\zeta(\alpha, t)|^2 d\alpha + \int_0^1 |\zeta_\alpha(\alpha, t)|^2 d\alpha.$$ (6.45)

**Proof:** We begin by proving exponential stability of the target system (6.33)–(6.35). Let $V(t)$ be the Lyapunov functional,

$$V(t) = \frac{m}{2} |w(0, t)|^2 + \frac{1}{2} |w(1, t)|^2 + \frac{1}{2} \int_0^1 |w_\alpha(\alpha, t)|^2 d\alpha,$$ (6.46)
where \( m \) is a positive scalar to be determined. In the sequel, we omit the arguments \((\alpha, t)\) unless needed for clarity. Taking the time derivative of \( V(t) \) gives

\[
\dot{V} = \frac{1}{2} \left[ mw_t(0)w(0)^* + mw(0)w_t(0)^* + w_t(1)w(1)^* \right.
\]

\[
+ w(1)w_t(1)^* + \int_0^1 (w_{\alpha t}^* w_{\alpha t} + w_{\alpha t}^* w_{\alpha t}) \, d\alpha \right],
\]

(6.47)

where \( w^* \) denotes the complex conjugate of \( w \). Substituting (6.35) yields

\[
\dot{V} = -cm|w(0)|^2 - c|w(1)|^2 + \text{Re} \left\{ \int_0^1 w_{\alpha t}^* w_{\alpha t} \, d\alpha \right\}.
\]

(6.48)

Integrating by parts and substituting (6.33)–(6.35) gives

\[
\dot{V} = -cm|w(0)|^2 - c|w(1)|^2 - \text{Re} \left\{ \int_0^1 \varepsilon w_{\alpha \alpha}^* w_{\alpha \alpha} \, d\alpha \right\},
\]

\[
- \varepsilon_R \int_0^1 |w_{\alpha \alpha}|^2 \, d\alpha + \text{Re} \left\{ w(0) \int_0^1 \varepsilon k_{\beta}(\alpha, 0) w_{\alpha \alpha}^* \, d\alpha \right\},
\]

\[
\leq -cm|w(0)|^2 - c|w(1)|^2 - c \int_0^1 |w_{\alpha}|^2 \, d\alpha
\]

\[
- \varepsilon_R \int_0^1 |w_{\alpha \alpha}|^2 \, d\alpha + |w(0)| \left| \int_0^1 \varepsilon k_{\beta}(\alpha, 0) w_{\alpha \alpha}^* \, d\alpha \right|.
\]

(6.49)

We now apply the Cauchy-Schwarz and Young’s inequalities with the parameter \( \rho > 0 \) to obtain

\[
\dot{V} \leq -cm|w(0)|^2 - c|w(1)|^2 - c \int_0^1 |w_{\alpha}|^2 \, d\alpha
\]

\[
- \varepsilon_R \int_0^1 |w_{\alpha \alpha}|^2 \, d\alpha + \frac{\rho}{2} |w(0)|^2 \int_0^1 |\varepsilon k_{\beta}(\alpha, 0)|^2 \, d\alpha + \frac{1}{2\rho} \int_0^1 |w_{\alpha \alpha}|^2 \, d\alpha,
\]

\[
= - \left( c - \frac{\rho \Lambda}{2m} \right) m|w(0)|^2 - c|w(1)|^2 - c \int_0^1 |w_{\alpha}|^2 \, d\alpha
\]

\[
- \left( \varepsilon_R - \frac{1}{2\rho} \right) \int_0^1 |w_{\alpha \alpha}|^2 \, d\alpha,
\]

(6.50)
where \( \Lambda = \int_0^1 |\varepsilon k_\beta(\alpha, 0)|^2 d\alpha \). Selecting the parameters, \( m = \rho \Lambda / c \) and \( \rho > 1/(2\varepsilon_R) \), we find

\[
\dot{V} \leq -c \frac{m}{2} |w(0)|^2 - c |w(1)|^2 - c \int_0^1 |w_\alpha|^2 d\alpha,
\]

\[
\leq -cV.
\]

(The choice of \( m \) assumes that \( c \neq b^2/(4\varepsilon) - \lambda \) so that \( \gamma \neq 0 \) and \( \Lambda > 0 \).)

From the Comparison Lemma [55] and Lemma C.1 in Appendix C, we have

\[
\Omega(t) \leq \frac{1}{p_1} \Psi(t) \leq \frac{1}{p_1 q_1} V(t) \leq \frac{1}{p_1 q_1} e^{-ct} V(0) \leq \frac{q_2}{p_1 q_1} e^{-ct} \Psi(0) \leq \frac{p_2 q_2}{p_1 q_1} e^{-ct} \Omega(0),
\]

where \( q_1 = \min\{1, 8\rho \Lambda / c\} / 16 \), \( q_2 = \max\{1, \rho \Lambda / c\} / 2 \), and \( p_1, p_2 \) are shown in (C.1), (C.2). The stability result is obtained from (6.52) with \( M = (p_2 q_2)/(p_1 q_1) \).

From Theorem 6.1, we see that the leader-enabled continuum design achieves exponential stability with a decay rate that can be arbitrarily set by the user, namely, the gain \( c \).

### 6.4 Leader Observer Design

In the previous sections, we have assumed that the leader agent has knowledge of all the agents’ positions. Since the leader agent is acting as a boundary actuator, it is also natural to use it as a boundary sensor, providing measurements to an observer that estimates the positions of all the agents. We now consider two scenarios: (A) the leader knows the position of itself, its nearest neighbor, and the anchor agent, and (B) the leader knows the position of only itself and its nearest neighbor. In both cases, the leader agent also knows the anchor’s bias term.

For these scenarios, we use backstepping for PDEs with boundary sensing [97] to design exponentially stable observers of the follower agents’ positions for use in the leader agent’s controller (6.40). As with boundary control, PDE backstepping leads to observer designs with closed-form observer gains, whereas other methods lead to more complicated designs.
6.4.1 Leader, Neighbor, and Anchor Measurements

Consider the observer for the PDE system (6.14)–(6.16),
\[
\dot{z}_t = \varepsilon \dot{z}_{\alpha\alpha} + b \dot{z}_\alpha + \lambda \dot{z} + p_1(\alpha) [z_\alpha(1) - \dot{z}_\alpha(1)] + p_2(\alpha) [z(1) - \dot{z}(1)],
\]
\[
\dot{z}_t(0) = -c \dot{z}(0) + c \bar{z}(0) + p_{00} [z(0) - \dot{z}(0)],
\]
\[
\dot{z}_t(1) = U_1(t) + p_{10} [z(1) - \dot{z}(1)],
\]
where \(\dot{z}(\alpha, t)\) is the position estimate of agent \(\alpha\) at time \(t\), \(U_1(t)\) is the leader agent’s control input, and \(z(0), z(1),\) and \(z_\alpha(1)\) are measurements. The functions, \(p_1(\alpha)\) and \(p_2(\alpha)\), and the constants, \(p_{00}\) and \(p_{10}\), are observer gains to be determined.

Define the observer error variable \(\tilde{z} = z - \dot{z}\) and consider the PDE error system,
\[
\tilde{z}_t = \varepsilon \tilde{z}_{\alpha\alpha} + b \tilde{z}_\alpha + \lambda \tilde{z} - p_1(\alpha) \tilde{z}_\alpha(1) - p_2(\alpha) \tilde{z}(1),
\]
\[
\tilde{z}_t(0) = -(c + p_{00}) \tilde{z}(0),
\]
\[
\tilde{z}_t(1) = -p_{10} \tilde{z}(1).
\]
We eliminate the advection term by substituting the change of variable \(\tilde{v}(\alpha) = \tilde{z}(\alpha)e^{\frac{b^2}{\varepsilon}(1-\alpha)}\) into (6.56)–(6.58) to obtain the reaction-diffusion equation
\[
\tilde{v}_t = \varepsilon \tilde{v}_{\alpha\alpha} + \left(\lambda - \frac{b^2}{4\varepsilon}\right) \tilde{v} - e^{-\frac{b^2}{\varepsilon}(1-\alpha)}p_1(\alpha) \tilde{v}_\alpha(1)
\]
\[
+ e^{-\frac{b^2}{\varepsilon}(1-\alpha)} \left(p_2(\alpha) - \frac{b}{2\varepsilon} p_1(\alpha)\right) \tilde{v}(1),
\]
\[
\tilde{v}_t(0) = -(c + p_{00}) \tilde{v}(0),
\]
\[
\tilde{v}_t(1) = -p_{10} \tilde{v}(1).
\]
Now define the coordinate transformation [97],
\[
\tilde{v}(\alpha) = \tilde{w}(\alpha) - \int_0^1 p(\alpha, \beta) \tilde{w}(\beta) \, d\beta,
\]
where the kernel \(p(\alpha, \beta)\) defined on \(\mathcal{S} = \{(\alpha, \beta) : 0 \leq \alpha \leq \beta \leq 1\}\) is given by
\[
p(\alpha, \beta) = -\delta \alpha \frac{I_1 \left( \sqrt{\delta (\beta^2 - \alpha^2)} \right)}{\sqrt{\delta (\beta^2 - \alpha^2)}},
\]
\[
\delta = \frac{1}{\varepsilon} \left( d + \lambda - \frac{b^2}{4\varepsilon} \right),
\]
\[
I_1 \left( \sqrt{\delta (\beta^2 - \alpha^2)} \right) = \int_0^1 \frac{e^{\frac{1}{\delta} \sqrt{\delta (\beta^2 - \alpha^2)}}}{\sqrt{\delta (\beta^2 - \alpha^2)}} \, d\beta.
\]
with \( d > 0 \). Using (6.62), we transform the error system (6.59)–(6.61) to the target system,

\[
\ddot{w}_t = \varepsilon \dot{w}_{\alpha \alpha} - d \ddot{w},  \\
\dot{w}_t(0) = -d\dot{w}(0),  \\
\dot{w}_t(1) = -d\dot{w}(1),
\]

(6.63)\((6.64)\)(6.65)

to determine the observer gains. From (6.59), (6.62), and (6.63), we find

\[
p_1(\alpha) = -\varepsilon \delta \alpha \frac{I_1\left(\sqrt{\delta (1 - \alpha^2)}\right)}{\sqrt{\delta (1 - \alpha^2)}} e^{\frac{b}{2}(1-\alpha)},
\]

(6.66)

\[
p_2(\alpha) = \varepsilon \delta^2 \alpha \frac{I_2\left(\sqrt{\delta (1 - \alpha^2)}\right)}{\delta (1 - \alpha^2)} e^{\frac{b}{2}(1-\alpha)} - \frac{(b + \varepsilon \delta) \delta \alpha}{2} \frac{I_1\left(\sqrt{\delta (1 - \alpha^2)}\right)}{\sqrt{\delta (1 - \alpha^2)}} e^{\frac{b}{2}(1-\alpha)},
\]

(6.67)

and from (6.60), (6.61), (6.62), (6.64), and (6.65), we have \( p_{00} = d - c \) and \( p_{10} = d \).

**Theorem 6.2** The observer (6.53)–(6.55) with gains (6.66), (6.67), \( p_{00} = d - c \), and \( p_{10} = d \) converges exponentially in the \( H^1 \) norm to the state \( z(\alpha, t) \), i.e., there exists \( N > 0 \) such that for all \( t \geq 0 \),

\[
\Sigma(t) \leq Ne^{-2dt} \Sigma(0),
\]

(6.68)

where \( c, d > 0 \) and

\[
\Sigma(t) = |\ddot{z}(0, t)|^2 + |\ddot{z}(1, t)|^2 + \int_0^1 |\ddot{z}(\alpha, t)|^2 d\alpha + \int_0^1 |\ddot{z}_\alpha(\alpha, t)|^2 d\alpha.
\]

(6.69)

**Proof:** The transformation (6.62) can be inverted to obtain, \( \dot{w}(\alpha, t) = \ddot{v}(\alpha, t) + \int_{\alpha}^{1} q(\alpha, \beta) \ddot{v}(\beta, t) \, d\beta \), where the inverse gain kernel \( q \in C^2(S) \) [96], so we begin by proving exponential stability of the target system (6.63)–(6.65). Let \( W(t) \) be the Lyapunov functional,

\[
W(t) = \frac{1}{2} |\ddot{w}(0, t)|^2 + \frac{1}{2} |\ddot{w}(1, t)|^2 + \frac{1}{2} \int_0^1 |\ddot{w}_\alpha(\alpha, t)|^2 d\alpha.
\]

(6.70)

In the sequel, we omit the arguments \((\alpha, t)\) unless needed for clarity.
Computing the time derivative of $W(t)$, integrating by parts, and substituting (6.63)–(6.65) yields

$$\dot{W} = -d|\dot{w}(0)|^2 - d|\dot{w}(1)|^2 - \text{Re}\left\{d\, w_w^*\big|_0^1 + \int_0^1 (\varepsilon\tilde{w}_{\alpha \alpha} - d\dot{w})\tilde{w}_{\alpha \alpha}^* \, d\alpha\right\}. \quad (6.71)$$

Integrating by parts gives

$$\dot{W} = -d\left(|\dot{w}(0)|^2 + |\dot{w}(1)|^2 + \int_0^1 |\tilde{w}_\alpha|^2 \, d\alpha\right) - \varepsilon_R \int_0^1 |\tilde{w}_{\alpha \alpha}|^2 \, d\alpha,$$

$$\leq -2dW. \quad (6.72)$$

From the Comparison Lemma [55] and Lemma C.2 in Appendix C, we have

$$\Sigma(t) \leq \frac{1}{r_1} \Phi(t) \leq \frac{16}{r_1} W(t) \leq \frac{16}{r_1} e^{-2dt} W(0) \leq \frac{8}{r_1} e^{-2dt} \Phi(0) \leq \frac{8r_2}{r_1} e^{-2dt} \Sigma(0), \quad (6.73)$$

where $r_1, r_2$ are shown in (C.5), (C.6). The stability result is obtained from (6.73) with $N = 8r_2/r_1$.

**6.4.2 Leader and Neighbor Measurements**

From $p_{00} = d - c$, we see that choosing the target system (6.63)–(6.65) with $d = c$ results in the observer gain, $p_{00} = 0$, nullifying the anchor agent position measurement. Thus, the transformation (6.62) transforms the error system (6.59)–(6.61) to the target system, $\tilde{w}_t = \varepsilon\tilde{w}_{\alpha \alpha} - c\tilde{w}$, with boundary conditions, $\tilde{w}_t(0) = -c\tilde{w}(0)$, $\tilde{w}_t(1) = -c\tilde{w}(1)$, and yields an observer that requires communication/sensing of only the leader agent’s nearest neighbor,

$$\dot{z}_t = \dot{z}_{\alpha \alpha} + b\dot{z}_\alpha + \lambda\dot{z} + p_1(\alpha) [z_\alpha(1) - \dot{z}_\alpha(1)] + p_2(\alpha) [z(1) - \dot{z}(1)], \quad (6.74)$$

$$\dot{z}_t(0) = -c\dot{z}(0) + c\dot{z}(0), \quad (6.75)$$

$$\dot{z}_t(1) = U_1(t) + c [z(1) - \dot{z}(1)], \quad (6.76)$$

where the observer gains, $p_1(\alpha)$ and $p_2(\alpha)$, are given by (6.66) and (6.67) by setting $\delta = \gamma$. From Theorem 6.2, we have exponential stability in the $H^1$ norm; specifically, $\Sigma(t) \leq Ne^{-2dt}\Sigma(0)$. Note how the decay rate is tied to the control gain $c$ and hence, cannot be arbitrarily set without knowledge of the anchor agent’s position.
6.4.3 Output Feedback

Equipped with the stabilizing feedback (6.40) and observer (6.53)–(6.55), one can now pursue output feedback for the leader agent to stabilize the deployment profile, namely, let \( U_1(t) = -K\{\hat{z}\}(t) + K\{\ddot{z}\} \) be the control input for the leader agent. To prove stability with output feedback, we must prove stability of the observer and observer-error system \((\hat{z}, \tilde{z})\), which can be converted to the target system \((\hat{w}, \tilde{w})\) using the change of variables, \(\hat{v}(\alpha) = (\hat{z}(\alpha) - \bar{z}(\alpha))e^{\frac{\epsilon \alpha}{\beta}}\) and \(\tilde{v}(\alpha) = \bar{z}(\alpha)e^{\frac{\epsilon \alpha}{\beta}}\), and the transformations, \(\hat{w}(\alpha) = \hat{v}(\alpha) - \int_0^\alpha k(\alpha, \beta)\tilde{v}(\beta) d\alpha\) and (6.62). The target system \((\hat{w}, \tilde{w})\) is

\[
\begin{align*}
\hat{w}_t &= \epsilon \hat{w}_{\alpha\alpha} - c\hat{w} - \epsilon k_\beta(\alpha, 0)\tilde{w}(0) + \epsilon (r(\alpha) - p_\beta(\alpha, 1))\tilde{w}(1) \\
&\quad + \epsilon (s(\alpha) + p(\alpha, 1))\tilde{w}(1) + \epsilon s(1)\tilde{w}(1) \\
\hat{w}_t(0) &= -c\hat{w}(0) + (d - c)\tilde{w}(0), \\
\hat{w}_t(1) &= -c\hat{w}(1) + (d + \epsilon r(1))\tilde{w}(1) + \epsilon s(1)\tilde{w}(1),
\end{align*}
\]

where the \(\tilde{w}\)-system is given by (6.63)–(6.65), \(r(\alpha) = \int_0^\alpha k(\alpha, \beta)p_\beta(\beta, 1) d\beta\), and \(s(\alpha) = -\int_0^\alpha k(\alpha, \beta)p(\beta, 1) d\beta\). This system is a cascade connection where the exponentially stable \(\tilde{w}\)-system drives the \(\hat{w}\)-system, which is also exponentially stable when unforced. Consequently, one would expect stability of this system to be in the \(H^1\) or higher norm.

6.5 Discretized Agent Control Laws

Implementable control laws for a finite number of agents are obtained by spatially discretizing the PDE model (6.14), the leader agent’s controller (6.40), and if employed, the observer (6.53)–(6.55). The anchor agent’s controller (6.37) does not require modification since it utilizes only the anchor’s own position. When we spatially discretize the system, the state variable \(z(\alpha, t)\) becomes \(z(ih, t)\) where \(i \in \{0, \ldots, n\}\), \(h = 1/n\), and \((n + 1)\) is the total number of agents. We denote the position of the anchor, follower, and leader agents as \(z_0, z_i\), and \(z_n\).

The discretization process gives rise to two questions: 1) “How many agents must be deployed, i.e., how fine must the spatial discretization be to retain stability
in the closed-loop system?” and 2) “Is the equilibrium of the spatially discretized system the same as the equilibrium of the continuous system?” One can address the first question by transforming the spatially discretized system using the transformation, $w(ih, t) = u(ih, t) - h \sum_{j=0}^{i-1} k(ih, jh) u(jh, t)$ where $k$ is the gain kernel (6.32), and performing a Lyapunov analysis on the resulting target system. Due to space limitations, we do not address this issue in the present paper, but similar results can be found in [6], [7], and [98]. Moreover, when compared to other design options, the benefits of this multi-agent control design, which allows the user to deploy large numbers of agents into various planar formations by communicating with only two agents, are mitigated when deploying few agents. Concerning the second question, we present a method that minimizes the approximation error due to the spatial discretization at the desired deployment $\bar{z}(ih)$. Hence, the achieved deployment $\bar{z}_i$ closely approximates $\bar{z}(ih)$. Note, however, that both questions 1) and 2) become irrelevant as the number of agents deployed is increased.

Using three-point central differencing to approximate the spatial derivatives in (6.14), we obtain the discretized follower agent control laws,

$$U_i(t) = \varepsilon \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} + b \frac{z_{i+1} - z_{i-1}}{2h} + \lambda z_i. \quad (6.80)$$

For the leader agent, we define the operator $K_D\{\cdot\}$, a discretized version of (6.38), acting on the vector $[\xi_0, \cdots, \xi_n]^T$, where $\xi_i(t) = \xi(ih, t)$, as

$$K_D\{\xi\}(t) = \varepsilon \sqrt{\gamma} I_1(\sqrt{\gamma}) e^{-\frac{\varepsilon}{4}} \xi_0(t) + \left( c + \frac{\varepsilon \gamma^2}{8} - \frac{\varepsilon \gamma}{2} + \frac{b \gamma}{4} \right) \xi_n(t)$$

$$+ \varepsilon \frac{\gamma}{2} \left( \frac{\xi_n(t) - \xi_{n-1}(t)}{h} \right) + \varepsilon \frac{\gamma h}{2} \left( f_n(t) + 2 \sum_{i=1}^{n-1} f_i(t) \right), \quad (6.81)$$

where $f_i(t) = e^{-\frac{\varepsilon}{2}(1-ih)} \Pi(ih)\xi_i(t)$, $f_n(t) = \left( \frac{\gamma}{4h} + \frac{1}{8} \right) \xi_n(t)$, and $\Pi$ is shown in (6.39). To obtain (6.81), we use the trapezoidal rule and the two-point backward difference to approximate the integral and $z_{a(1)}$ terms in (6.38). The leader’s discretized control law is simply,

$$U_n(t) = -K_D\{z\}(t) + K_D\{\bar{z}\}. \quad (6.82)$$

Together, the control laws (6.37), (6.80), and (6.82) govern the dynamics of the
agents according to the linear ODE system,

\[
\dot{z}_0 = U_0(t), \quad \dot{z}_i = U_i(t), \quad \dot{z}_n = U_n(t).
\]

(6.83)

We see directly from (6.80) that the spatial discretization imposes a fixed communication topology on the follower agents, specifically a chain graph, since the followers depend only on their nearest neighbors on the graph. The leader agent requires global information to stabilize the deployment as seen by (6.82). However, the observers designed in Section 6.4 can be discretized in a similar manner so that if output feedback were employed, the leader agent’s control law is

\[
U_n(t) = -\mathcal{K}_D\{\tilde{z}\}(t) + \mathcal{K}_D\{\bar{z}\},
\]

(6.84)

which requires only local information.

Due to the approximation error inherent in the spatial discretization, the equilibrium of (6.83), \( \bar{z}_i \), will not equal the desired deployment, \( \bar{z}(ih) \). Evaluating (6.83) at \( z_i = \bar{z}(ih) \), we have \( U_0(t) = -c\bar{z}(0) + c\bar{z}(0) = 0 \), \( U_n(t) = -\mathcal{K}_D\{\tilde{z}\}(t) + \mathcal{K}_D\{\bar{z}\} = 0 \), and

\[
U_i(t) = \varepsilon \frac{\bar{z}_{i+1} - 2\bar{z}_i + \bar{z}_{i-1}}{h^2} + \beta \frac{\bar{z}_{i+1} - \bar{z}_{i-1}}{2h} + \lambda \bar{z}_i,
\]

(6.85)

where \( O(h^2) \) is the accuracy of the three-point central difference [84]. Because the desired deployment is encoded in the leader and anchor agents’ bias terms, \( U_0(t) = U_n(t) = 0 \) when \( z_i = \bar{z}(ih) \). The followers, however, are unaware of \( \bar{z}(ih) \), and at \( z_i = \bar{z}(ih) \), \( U_i(t) = O(h^2) \neq 0 \).

To make \( \bar{z}_i \) closely approximate \( \bar{z}(ih) \), we need to minimize the approximation error \( O(h^2) \), but we must do so for the entire deployment family, not for specific deployments. Otherwise, the error minimization will adversely affect the flexibility of the design and will require the follower agents to know the desired deployment. One method is to simply deploy more agents since \( O(h^2) \rightarrow 0 \) as \( n \rightarrow \infty \), but we seek a method to decrease the error for a fixed \( n \). We exploit our knowledge of the basis functions (6.18) that comprise the desired deployment.
family, allowing us to characterize a measure of the approximation error for the entire family.

Define vectors for the family of basis function spatial derivatives as

\[ \Psi_{\alpha\alpha} = [\psi_{0,\alpha\alpha}(h), \ldots, \psi_{0,\alpha\alpha}(nh-h), \ldots, \psi_{3,\alpha\alpha}(h), \ldots, \psi_{3,\alpha\alpha}(nh-h)]^T, \]  
\[ \Psi_{\alpha} = [\psi_{0,\alpha}(h), \ldots, \psi_{0,\alpha}(nh-h), \ldots, \psi_{3,\alpha}(h), \ldots, \psi_{3,\alpha}(nh-h)]^T \]

where \( \psi_{k,\alpha\alpha}(ih) \) and \( \psi_{k,\alpha}(ih) \) are the second and first spatial derivatives of the basis function \( \psi_k \) for agent \( i, i \in \{1, \ldots, n-1\} \). The vectors for the approximations of the spatial derivatives, \( \hat{\Psi}_{\alpha\alpha} \) and \( \hat{\Psi}_{\alpha} \), are defined similarly but with the three-point central difference approximations,

\[ \hat{\psi}_{k,\alpha\alpha}(ih) = \frac{1}{h_2} (\psi_k(ih+h) - 2\psi_k(h) + \psi_k(ih-h)), \]  
\[ \hat{\psi}_{k,\alpha}(ih) = \frac{1}{2h} (\psi_k(ih+h) - \psi_k(ih-h)) \]

The actual derivatives and approximations are related by \( \psi_{k,\alpha\alpha}(ih) = \hat{\psi}_{k,\alpha\alpha}(ih) + O(h^2) \) and \( \psi_{k,\alpha}(ih) = \hat{\psi}_{k,\alpha}(ih) + O(h^2) \).

Now, define the second and first spatial derivative approximation error for the family of basis functions as \( \tilde{\Psi}_{\alpha\alpha} = \Psi_{\alpha\alpha} - \hat{\Psi}_{\alpha\alpha} \), and \( \tilde{\Psi}_{\alpha} = \Psi_{\alpha} - \hat{\Psi}_{\alpha} \), and characterize the error as

\[ \| \tilde{\Psi}_{\alpha\alpha} \|_2 = \mu_0 \leq \tilde{\mu}_0, \quad \| \tilde{\Psi}_{\alpha} \|_2 = \mu_1 \leq \tilde{\mu}_1, \]

where \( \tilde{\mu}_0 \) and \( \tilde{\mu}_1 \) represent user-determined thresholds for the approximation error and \( \| \cdot \|_2 \) denotes the Euclidean norm. If both inequalities are satisfied, the achieved deployment \( \bar{z}_i \) will match the desired deployment \( \bar{z}(ih) \) according to the user’s specifications.

But what if one or both of the above inequalities are not satisfied? We propose the modified follower agent control law,

\[ U_i(t) = \bar{\xi}z_{i+1} - \frac{2z_i + z_{i-1}}{h} + \frac{\bar{\xi}z_{i+1} - z_{i-1}}{2h} + \lambda z_i, \]

where \( \bar{\xi} = \varepsilon(1 + \nu_0) \), \( \bar{b} = b(1 + \nu_1) \), \( \nu_0 = \arg \min_\nu f_0(\nu) \), \( f_0(\nu) = \| \hat{\Psi}_{\alpha\alpha} \nu - \tilde{\Psi}_{\alpha\alpha} \|_2^2 \), \( \nu_1 = \arg \min_\nu f_1(\nu) \), and \( f_1(\nu) = \| \hat{\Psi}_{\alpha} \nu - \tilde{\Psi}_{\alpha} \|_2^2 \). Hence, (6.91) is minimal in a least squares sense at \( z_i = \bar{z}(ih) \) because the numerical approximation error
for the family of basis function spatial derivatives has been minimized by using the approximations $(1 + \nu_0)\hat{\psi}_{k,\alpha\alpha}(ih), (1 + \nu_1)\hat{\psi}_{k,\alpha}(ih)$ for the actual derivatives $\psi_{k,\alpha\alpha}(ih), \psi_{k,\alpha}(ih)$. Since $f_0(\nu)$ and $f_1(\nu)$ are minimized over a single parameter, the follower agents can still achieve a family of deployments while employing a single set of gains (6.91). For 1-D deployments, the procedure to find $\nu_0$ and $\nu_1$ is the same, but the minimization is done using two basis functions instead of four.

If $\mu_0 \leq \bar{\mu}_0$ and/or $\mu_1 \leq \bar{\mu}_1$ is satisfied, $\nu_0$ and/or $\nu_1$ can be simply set to zero. For example, the three-point central difference is exact for linear deployments, so $\nu_0 = \nu_1 = 0$. If after computing $\nu_0$ and/or $\nu_1$, the inequalities $f_0(\nu_0) \leq \bar{\mu}_0$ and/or $f_1(\nu_1) \leq \bar{\mu}_1$ are not satisfied, improved deployment accuracy can be achieved only by deploying more agents. As an example, consider a 1-D deployment modeled by $x_t = x_{\alpha\alpha} + 4\pi^2x$, which has basis functions, $(\cos(2\pi\alpha), \sin(2\pi\alpha))$. For deployments with 10 agents ($n = 9$), the approximation error is $\|\tilde{\Psi}_{\alpha\alpha}\|_2 = 4.4622$ when the followers employ (6.80) and is reduced to $f_0(\nu_0) = 1.2841 \times 10^{-13}$, where $\nu_0 = 0.0416$, when implementing (6.91) instead, indicating that the achieved deployment $\bar{z}_i$ is a much closer approximation of $\bar{z}(ih)$ when using the feedback laws (6.91).

While increasing the number of agents causes $\bar{z}_i$ to better approximate $\bar{z}(ih)$ because the discretization step size $h$ decreases, this smaller $h$ leads to large gains in the follower agents feedback laws (6.91). However, we can simply scale all the control laws to prevent disproportionately large gains and hence, velocity commands. Specifically, we implement the system, $\dot{z}_0 = \kappa U_0(t), \dot{z}_i = \kappa U_i(t), \dot{z}_n = \kappa U_n(t)$, which has a decay rate of $\kappa c$ instead of $c$. This scaling is meant for the implementation of the control strategies on velocity-actuated systems. Alternatively, this deployment paradigm could be used as a planning algorithm where $z(\alpha, t)$ represents a non-physical, consensus-like variable such that no physical actuation is required, making this scaling unnecessary.

### 6.6 Simulations

We present a variety of deployment examples to demonstrate the flexibility of our results. All deployments, unless otherwise stated, are simulated with 10
Figure 6.3: Agents in the ellipse deployment family (a) rendezvous at the origin with (b) the corresponding control effort (top) and observer error (bottom). The agents converge to the origin after the observer error is negligible ($\approx 14$ sec). In (b), we plot the control effort and observer error for $t \in [0,15]$ sec to highlight the transients since all the curves decay to zero after 15 sec.

agents ($n = 9$) where the anchor and follower agents employ (6.37) and (6.91), and the leader agent employs output feedback (6.84) using the observer (6.53)–(6.55) with $d = 10$. The agents use the scaling $\kappa = h/4$, and the leader and anchor agents select the control gain $c = 7.5$. We provide the parameter set ($\varepsilon, b, \lambda, \nu_0, \nu_1$) for each PDE model used. In Figs. 6.4–6.6, the trajectories are shaded from light to dark as the agents move from one deployment to another with the desired planar curves shaded green. The desired deployments are also projected onto the $x$-$y$ plane when needed for clarity.

Our first example focuses on the family of elliptical deployments that are centered about the origin, which is derived from identical, unstable reaction-diffusion equations $(1,0,4\pi^2,0.0416,0)$. In Figure 6.3(a), the agents’ initial positions are sampled from the Gaussian distribution $\mathcal{N}(0,1)$ (zero mean with unity variance) and rendezvous at the origin due to the movements of the anchor and the leader agents, whose bias terms are set to zero. The observer’s initial condition is $(\hat{x}_i(0), \hat{y}_i(0)) = (x_i(0) + \xi_i, y_i(0) + \eta_i), i \in \{0, \ldots, n-1\}$, where $(\xi_i, \eta_i)$ is sampled from the distribution $\mathcal{N}(0,0.04)$, and $(\hat{x}_n(0), \hat{y}_n(0)) = (x_n(0), y_n(0))$.
Figure 6.3(b) shows the control effort $\| [u_i, v_i] \|_2$ and the norm of the observer error $\| [x_i - \hat{x}_i, y_i - \hat{y}_i] \|_2$ for each agent. Note how the agents exhibit a transient before converging toward the origin at approximately 14 sec, which is when the observer error becomes negligible. In Figure 6.4(a), the anchor and leader change their bias terms to deploy the agents onto a circle

$$
\bar{x}(\alpha) = 2 \cos(2\pi \alpha),
$$

$$
\bar{y}(\alpha) = 2 \sin(2\pi \alpha),
$$

and then onto an ellipse rotated counterclockwise about the origin by $\phi = \pi/6$ rad,

$$
\bar{x}(\alpha) = \frac{5\sqrt{3}}{2} \cos(2\pi \alpha) - \frac{3}{2} \sin(2\pi \alpha),
$$

$$
\bar{y}(\alpha) = \frac{5}{2} \cos(2\pi \alpha) + \frac{3\sqrt{3}}{2} \sin(2\pi \alpha).
$$

For the following examples, we could again initialize the agents randomly in the plane with an initial observer error, but instead we assume convergence of the observer and initialize the agents at the origin to depict the agents deploying, and then redeploying, more clearly. Figure 6.4(b) depicts the agents deploying onto a
**Figure 6.5**: Agent trajectories (a) using an advection-diffusion equation and a reaction-diffusion equation for the $x$- and $y$-deployments respectively, and (b) for the complex-valued circle deployment family when deploying to a circle centered at $(0, 0)$ and to another centered at $(4, -4)$.

figure 8,

\[
\bar{x}(\alpha) = 2 \sin(2\pi \alpha), \\
\bar{y}(\alpha) = 2 \sin(4\pi \alpha),
\]

and a distorted figure 8,

\[
\bar{x}(\alpha) = 3 \cos(2\pi \alpha) + 6 \sin(2\pi \alpha), \\
\bar{y}(\alpha) = 6 \sin(4\pi \alpha),
\]

which are in, what we term, the Lissajous figure 8 family. This family is modeled using two unstable reaction-diffusion equations where the parameters $(1, 0, 4\pi^2, 0.0416, 0)$ model the $x$-axis deployment and $(1, 0, 16\pi^2, 0.1796, 0)$, the $y$-axis deployment. The deployments depicted in Figure 6.5(a) are formed using the advection-diffusion equation $(1, 5, 0, -0.0253, -0.0496)$ for the $x$-axis deployment and an unstable reaction-diffusion equation $(1, 0, 2\pi^2, 0.0206, 0)$ for the $y$-axis deployment. At $t = 30$ sec, the agents stabilize

\[
\bar{x}(\alpha) = \frac{2}{e^{-5} - 1} \left( -e^{-5} - 1 + 2e^{-5\alpha} \right), \\
\bar{y}(\alpha) = 3 \sin(\sqrt{2\pi} \alpha)
\]
Figure 6.6: Agent trajectories for (a) the complex-valued spiral deployment family when deploying to a logarithmic spiral, and (b) a helical deployment with \( t \in [0, 30] \) sec.

Before redeploying to
\[
\bar{x}(\alpha) = \frac{5}{e^{-5} - 1} \left( -e^{-5} - 1 + 2e^{-5\alpha} \right),
\]
\[
\bar{y}(\alpha) = 3 \cos(\sqrt{2}\pi\alpha) + 4 \sin(\sqrt{2}\pi\alpha).
\]

We now provide some examples that utilize Ginzburg-Landau PDE models. In Figure 6.5(b), the agents deploy to a circle centered at \((0, 0)\),
\[
\bar{x}(\alpha) = -\frac{5}{2} \sin(2\pi\alpha),
\]
\[
\bar{y}(\alpha) = \frac{5}{2} \cos(2\pi\alpha),
\]
and redeploy to a larger circle centered at \((4, -4)\),
\[
\bar{x}(\alpha) = 4 - 5 \sin(2\pi\alpha),
\]
\[
\bar{y}(\alpha) = -4 + 5 \cos(2\pi\alpha),
\]
using the complex-valued reaction-advection equation \((1/2 + j/2, \pi - j\pi, 0, 0.0416, 0.0861)\), which generates circles centered about \((a_0, a_1)\). The ability to move the circles about the plane makes this deployment applicable to capture and escort missions [51, 80]. A logarithmic spiral,
\[
\bar{x}(\alpha) + j\bar{y}(\alpha) = -2(\sin(6.28\alpha) + j \cos(6.28\alpha))e^{-2.49\alpha},
\]
is stabilized in Figure 6.6(a) by using a complex-valued reaction-advection-diffusion equation \((1 - j, 25/4 - j3/4, 9\pi^2/4 - j11\pi^2/4, 0.301, 0.486)\). For this deployment family, the desired deployment is not precisely reached due to the approximation error incurred by deploying 10 agents (whereas the deployment error in the other examples is negligible for 10 agents). This error is significantly reduced by deploying 15 agents and is indiscernible when deploying 30 agents. The final leader-enabled example (Figure 6.6(b)) is a 3-D deployment to a helix,

\[
\begin{align*}
\bar{x}(\alpha) &= -2\sin(4\pi\alpha), \\
\bar{y}(\alpha) &= 2\cos(4\pi\alpha), \\
\bar{z}(\alpha) &= -10 + 20\alpha,
\end{align*}
\]

which is created by combining a deployment derived from a complex-valued advection-diffusion equation \((1/2 + j/2, 2\pi - j2\pi, 0, 0.699, 0.1481)\) with a linear deployment generated by the 1-D heat equation.

Requiring a leader agent to achieve these deployments does give rise to questions about the method’s robustness regarding deployment stabilization when faced with agent failures. We categorize a failure as an agent that stops moving prematurely, which leads to this failed position being continually used in the feedback laws of the failed agent’s neighbors. Consequently, the failure of an agent can cause an unstable collective response when employing an open-loop unstable PDE model. If the failed agent is the anchor, however, its failure has the same effect as selecting a different value for its bias term, causing the agents to stabilize another (unintended) deployment. If the PDE model is open-loop stable, the agents achieve a stable, but again unintended, deployment no matter which agent fails since a failed agent acts like an anchor, splitting the deployment into two formations—one between the anchor and the failed agent and another between the failed agent and the leader. In short, the failure of any agent may prevent the agents from achieving the desired deployment, but an unstable response is possible only when an open-loop unstable PDE model is employed.

If the agents have the ability to detect failed agents (that are not the leader), remove them from the topology, and reconnect the topology, the agents would
stabilize a deployment that differs slightly from the desired one since the agents’ gains are based on the initial number of deployed agents. If the leader were to fail, the agents could elect a new leader, but the user would have to provide this new leader (a former follower agent) with the necessary desired deployment information.

Another concern is collision avoidance. To prevent collisions, potential field-based controllers (or other suitable controllers) could be appended to the controllers used for deployment, but modifications would be needed for desired deployments that have collocated agents. As mentioned at the end of Section 6.5, this methodology can also be used as a planning algorithm where \( z(\alpha, t) \) is a non-physical variable that converges to the desired deployment \( \bar{z}(\alpha) \) while the agents are stationary. Once \( z(\alpha, t) \) has converged, each agent effectively knows its desired position and may deploy to this position using any controller that prevents collisions.

### 6.7 Summary

We have introduced a PDE-based approach for the deployment of agents onto families of planar curves. The PDE models are a spatial application of the internal model principle, meaning they make a deployment family feasible but do not guarantee stability. While the standard diffusion-based feedback leads to inherently stable deployment to linear formations, for which leader assistance is not needed, the nonlinear (in \( \alpha \)) curves that we pursue may be open-loop unstable. We employ a leader and anchor agent, whose controllers are designed using PDE boundary control techniques, to select and stabilize a desired deployment from the feasible deployments.

It is crucial to observe that in our framework the follower agents deploy out and maintain a stable formation not by being commanded some reference positions, but by being induced (indirectly influenced) by the leader’s and anchor’s motions. Such a paradigm allows a user to control the formation geometry of an entire group of agents by adjusting the bias terms of only two agents. With a leader agent, the agents can stably deploy to the rich geometric family of formations derived from the reaction-advection-diffusion PDE class, which includes
several deployments of practical interest—for example, the leader deploying out the majority of agents near a target position (occupied by the anchor) while staying at base, or conversely, deploying out only a few agents; deploying agents out in both directions symmetrically while both the leader and anchor stay at base, creating a “protective shell” for the leader; or deploying the agents to encircle a point of interest for surveillance or escort purposes.

If both the boundary agents are treated as virtual agents, more traditional Dirichlet and Neumann boundary conditions can be used, which may be necessary to extend this work to the stabilization of deployments governed by nonlinear PDE models such as the Burgers equation, which allows for shock-like deployments. Extending this paradigm to planar surface deployments and performing motion planning to control the agents’ transient behavior when deploying are also of interest.


The dissertation author is the primary investigator and author of this work.
Chapter 7

Multi-agent Deployment to a Family of Planar Arcs

We introduce a nonlinear PDE model to achieve stable deployment to circular arcs with either uniform or nonuniform agent spacing. This PDE model incorporates the follower agents’ control laws, which employ a spatial internal model principle to allow the agents to converge to circular arc formations of radius $R$. These geometric curves are formed using controllers that utilize only local information, are parameterized in $R$, and do not require further knowledge of the desired deployment. The control laws for the discrete agents are obtained by spatially discretizing the PDE model. Figure 7.1 depicts the communication topology imposed by the spatial discretization.

Boundary agents are used to manipulate the geometry of the deployment through the boundary conditions, as in [57]. We refer to both boundary agents as “anchor” agents rather than “leader” agents, according to the terminology presented in Chapter 6 ([37]), since this model does not require leader feedback for stabilization but does utilize boundary agents to influence the followers. By moving the anchor agents about the circle, the agents achieve different deployments without changing the followers’ gains. We extend the circular arc deployment family to include elliptical arcs and spirals by replacing the parameter $R$ with the variable $R(\alpha, t)$, which creates a dynamic control law for the follower agents. By setting boundary values for $R(\alpha, t)$, the anchor agents can manipulate the radial
deployment of the followers as well as their angular position. We illustrate our results with numerical simulations.

This chapter is organized as follows: Section 7.1 introduces the nonlinear PDE model and presents the convergence result, Section 7.2 extends the model to include elliptical arcs and spirals, Section 7.3 discretizes the infinite-dimensional controllers for \( N \) agents, Section 7.4 presents numerical simulations, and Section 7.5 summarizes the results.

### 7.1 Deployment to Circular Arcs

We seek to stably deploy agents to circular arc formations that can be rotated, stretched, and shortened about the circle with radius \( R \), while employing the same controller for each follower agent. We denote the position at time \( t \) of an agent indexed by \( \alpha \in [0, 1] \) in a large (continuum) group of fully actuated agents as \((x(\alpha, t), y(\alpha, t))\). The agents’ dynamics are modeled by

\[
\frac{\partial}{\partial t} x(\alpha, t) = U(\alpha, t),
\]

\[
\frac{\partial}{\partial t} y(\alpha, t) = V(\alpha, t),
\]
which, when discretized with respect to $\alpha$, leads to the following dynamical model for agent $i$:

\begin{align*}
\dot{x}_i(t) &= U_i(t), \\
\dot{y}_i(t) &= V_i(t).
\end{align*}

For later use, we define the notation: $x_{t}(\alpha,t) = \frac{\partial}{\partial t} x(\alpha,t)$, $x_{\alpha}(\alpha,t) = \frac{\partial}{\partial \alpha} x(\alpha,t)$, $x_{\alpha\alpha}(\alpha,t) = \frac{\partial^2}{\partial \alpha^2} x(\alpha,t)$.

### 7.1.1 Nonlinear PDE Model

Consider the nonlinear PDE model, which provides the velocity-actuated feedback laws for the follower agents ($0 < \alpha < 1$), where we have omitted the arguments $(\alpha,t)$ in the control laws for conciseness:

\begin{align*}
U(\alpha,t) &= -\frac{c}{2} x(x^2 + y^2 - R^2) - \varepsilon y \frac{y_{\alpha\alpha} - y_{x\alpha\alpha}}{x^2 + y^2} \\
&\quad - 2\varepsilon x \frac{y(x_{\alpha} - y_{\alpha}) - x_{\alpha} y_{\alpha}(x^2 - y^2)}{(x^2 + y^2)^2} - by \frac{y_{\alpha} - y_{x\alpha}}{x^2 + y^2}, \\
V(\alpha,t) &= -\frac{c}{2} y(x^2 + y^2 - R^2) + \varepsilon x \frac{y_{\alpha\alpha} - y_{x\alpha\alpha}}{x^2 + y^2} \\
&\quad + 2\varepsilon x \frac{y(x_{\alpha} - y_{\alpha}) - x_{\alpha} y_{\alpha}(x^2 - y^2)}{(x^2 + y^2)^2} + bx \frac{y_{\alpha} - y_{x\alpha}}{x^2 + y^2},
\end{align*}

where $c, \varepsilon > 0$, $b$ is a scalar coefficient, and $R > 0$ is the desired radius of the circular arc formations.

To manipulate the geometry of the formation, we employ two anchor agents, one at each boundary ($\alpha = 0$ and $\alpha = 1$). We define their control laws, which serve as boundary conditions for the PDE model (7.1)–(7.6), as

\begin{align*}
x_{t}(l,t) &= -\frac{c}{2} x(l,t) \left( x^2(l,t) + y^2(l,t) - R^2 \right) \\
&\quad + dy(l,t) \left( \arctan \frac{y(l,t)}{x(l,t)} - \arctan \frac{\bar{y}(l)}{\bar{x}(l)} \right), \\
y_{t}(l,t) &= -\frac{c}{2} y(l,t) \left( x^2(l,t) + y^2(l,t) - R^2 \right) \\
&\quad - dx(l,t) \left( \arctan \frac{y(l,t)}{x(l,t)} - \arctan \frac{\bar{y}(l)}{\bar{x}(l)} \right),
\end{align*}
where \( d > b^2/(4\varepsilon) \), \((\bar{x}(l), \bar{y}(l))\) is the desired position such that \( \bar{x}^2(l) + \bar{y}^2(l) = R^2 \), \( \arctan(\bar{y}(l)/\bar{x}(l)) = \theta_l \), and \( l = 0 \) or \( 1 \). With (7.7), (7.8), the anchor agents deploy to the circle with radius \( R \) at the given angles \( \theta_0 \) and \( \theta_1 \), respectively.

The control laws, (7.5), (7.6), cause the follower agents to deploy to a circular arc between the angles \( \theta_0 \) and \( \theta_1 \), provided no agent (including the anchors) are initialized at the origin. The agents are uniformly spaced about the arc if \( b = 0 \); otherwise, as \( \alpha \) increases, the agents become more dense when \( b > 0 \) and less dense when \( b < 0 \). This deployment is seen more readily by converting to polar coordinates, i.e.,

\[
\begin{align*}
  x(\alpha,t) &= r(\alpha,t) \cos \theta(\alpha,t), \\
  y(\alpha,t) &= r(\alpha,t) \sin \theta(\alpha,t),
\end{align*}
\]

where \( r(\alpha,t) = \sqrt{x^2(\alpha,t) + y^2(\alpha,t)} \) and \( \theta(\alpha,t) = \arctan(y(\alpha,t)/x(\alpha,t)) \), to obtain

\[
\begin{align*}
  U(\alpha,t) &= \frac{c}{2} r \cos \theta \left( R^2 - r^2 \right) - \varepsilon \left( r \sin \theta \right) \theta_{\alpha\alpha} - b \left( r \sin \theta \right) \theta_\alpha, \\
  V(\alpha,t) &= \frac{c}{2} r \sin \theta \left( R^2 - r^2 \right) + \varepsilon \left( r \cos \theta \right) \theta_{\alpha\alpha} + b \left( r \cos \theta \right) \theta_\alpha,
\end{align*}
\]

which results in the system,

\[
\begin{align*}
  (r^2)_t(\alpha,t) &= -cr^2(\alpha,t)(r^2(\alpha,t) - R^2), \\
  \theta_t(\alpha,t) &= \varepsilon \theta_{\alpha\alpha}(\alpha,t) + b \theta_\alpha(\alpha,t),
\end{align*}
\]

with boundary conditions,

\[
\begin{align*}
  (r^2)_t(l,t) &= -cr^2(l,t) \left( r^2(l,t) - R^2 \right), \\
  \theta_t(l,t) &= -d \left( \theta(l,t) - \theta_l \right).
\end{align*}
\]

It should be noted that the \( r^2 \)-subsystem is an \( \alpha \)-parameterized family of ODEs, whereas the \( \theta \)-subsystem is a PDE.

### 7.1.2 Convergence Result

Before stating a convergence result for the agents in the \( x-y \) plane, we need the following lemmas.
Lemma 7.1 For each $\alpha$, the $r^2$-subsystem (7.13), (7.15) has the solution

$$r^2(\alpha, t) = \frac{r_0^2(\alpha) R^2}{r_0^2(\alpha) - \tilde{r}_0^2(\alpha) e^{-cR^2 t}},$$

(7.17)

where $r_0(\alpha) = r(\alpha, 0)$, $\tilde{r}_0^2(\alpha) = r_0^2(\alpha) - R^2$. For every $r_0(\alpha) > 0$, $r(\alpha, t) \rightarrow R$ as $t \rightarrow \infty$.

Analyzing the $\theta$-subsystem in (7.14), (7.16) is more subtle. From the boundary conditions (7.16), we have

$$\theta(l, t) = e^{-dt} (\theta(l, 0) - \theta_l) + \theta_l,$$

(7.18)

so $\theta(0, t) \rightarrow \theta_0$ and $\theta(1, t) \rightarrow \theta_1$ as $t \rightarrow \infty$. Consequently the equilibrium profiles of (7.14) satisfy the two-point boundary value problem,

$$\varepsilon \bar{\theta}''(\alpha) + b \bar{\theta}'(\alpha) = 0,$$

(7.19)

$$\bar{\theta}(0) = \theta_0,$$

(7.20)

$$\bar{\theta}(1) = \theta_1,$$

(7.21)

whose solutions may be classified according to value of $b$.

In general, the equilibrium profile has the form, $\bar{\theta}(\alpha) = a_0 \psi_0(\alpha) + a_1 \psi_1(\alpha)$ where $a_0, a_1$ are scalars and $\psi_0, \psi_1$ are basis functions. If $b = 0$, the equilibrium is a linear profile, $\bar{\theta}(\alpha) = a_0 + a_1 \alpha$, and if $b \neq 0$, the equilibrium profile is $\bar{\theta}(\alpha) = a_0 + a_1 e^{-ab/\varepsilon}$. However, $\theta(\alpha, t) = \theta(\alpha, t) + 2j\pi$, $j \in \mathbb{N}$, so there exist countably many equilibria that lie “between” the final anchor agent positions $\bar{\theta}(0) = \theta_0$ and $\bar{\theta}(1) = \theta_1$. More specifically, if $b = 0$, we have

$$\bar{\theta}(\alpha) = (\theta_0 \pm 2j\pi) + (\theta_1 - \theta_0 \pm 2k\pi) \alpha,$$

(7.22)

and if $b \neq 0$,

$$\bar{\theta}(\alpha) = \frac{(\theta_0 \pm 2j\pi) e^{-\frac{b}{2\varepsilon}} - (\theta_1 \pm 2k\pi)}{e^{-\frac{b}{2\varepsilon}} - 1} + \frac{\theta_1 - \theta_0 \pm 2(k - j)\pi}{e^{-\frac{b}{2\varepsilon}} - 1} e^{-\frac{b}{2\varepsilon}},$$

(7.23)

where $j, k \in \mathbb{N}$.

The following lemma shows that the $\theta$-subsystem is exponentially stable in the $H^1$ norm, defined in (6.42).
Lemma 7.2 The $\theta$-subsystem (7.14), (7.16) is exponentially stable in the $H^1$ norm at the equilibrium, $\theta(\alpha,t) \equiv \bar{\theta}(\alpha)$, i.e., there exists $M > 0$ such that for all $t > 0$,

$$\Gamma(t) \leq Me^{-2\eta t}\Gamma(0), \quad (7.24)$$

where

$$\Gamma(t) = \bar{\theta}^2(0, t) + \bar{\theta}^2(1, t) + \int_0^1 \bar{\theta}^2(\alpha, t) d\alpha + \int_0^1 \bar{\theta}_\alpha^2(\alpha, t) d\alpha, \quad (7.25)$$

$$\bar{\theta}(\alpha, t) = \theta(\alpha, t) - \bar{\theta}(\alpha), \quad (7.26)$$

$$\eta = \min\left(\frac{\varepsilon}{8}, \frac{d}{2}\right) \left(1 - \frac{\mu |b|}{2d}\right), \quad (7.27)$$

$$d > \frac{b^2}{4\varepsilon}, \quad (7.28)$$

and $\varepsilon, \mu > 0$.

Proof: Working in the error variable (7.26), we obtain the error system,

$$\tilde{\theta}_t(\alpha, t) = \varepsilon \tilde{\theta}_{\alpha\alpha}(\alpha, t) + b\tilde{\theta}_\alpha, \quad (7.29)$$

$$\tilde{\theta}_t(0, t) = -d\tilde{\theta}(0, t), \quad (7.30)$$

$$\tilde{\theta}_t(1, t) = -d\tilde{\theta}(1, t). \quad (7.31)$$

Let $W(t)$ be the Lyapunov functional,

$$W(t) = \frac{m_0}{2} \bar{\theta}^2(0, t) + \frac{m_1}{2} \bar{\theta}^2(1, t) + \frac{1}{2} \int_0^1 \bar{\theta}^2(\alpha, t) d\alpha + \frac{\varepsilon}{2d} \int_0^1 \bar{\theta}_\alpha^2(\alpha, t) d\alpha, \quad (7.32)$$

where $m_0$ and $m_1$ are scalars to be determined. In the sequel, we omit the arguments $(\alpha, t)$ unless needed for clarity.

Taking the time derivative and substituting (7.29), (7.31) gives

$$\dot{W} = -dm_0 \bar{\theta}^2(0) - dm_1 \bar{\theta}^2(1) + \int_0^1 \bar{\theta} \left(\varepsilon \tilde{\theta}_{\alpha\alpha} + b\tilde{\theta}_\alpha\right) d\alpha + \frac{\varepsilon}{d} \int_0^1 \bar{\theta}_\alpha \tilde{\theta}_\alpha d\alpha. \quad (7.33)$$

Integrate by parts to obtain

$$\dot{W} = -dm_0 \bar{\theta}^2(0) - dm_1 \bar{\theta}^2(1) + \varepsilon \tilde{\theta}\tilde{\theta}_\alpha|_0^1 + \varepsilon \tilde{\theta}_\alpha\tilde{\theta}_t|_0^1$$

$$- \varepsilon \int_0^1 \bar{\theta}^2_\alpha d\alpha + b \int_0^1 \tilde{\theta}_\alpha \tilde{\theta}_\alpha d\alpha - \frac{\varepsilon}{d} \int_0^1 \bar{\theta}_{\alpha\alpha} \tilde{\theta}_t d\alpha. \quad (7.34)$$
Substitute (7.29), (7.31), integrate the second term on the second line of (7.34), and bound terms to yield
\[
\dot{W} \leq -dm_0 \tilde{\theta}^2(0) - dm_1 \tilde{\theta}^2(1) + \frac{|b|}{2} \tilde{\theta}^2(0) + \frac{|b|}{2} \tilde{\theta}^2(1) \\
- \varepsilon \int_0^1 \tilde{\theta}_a^2 d\alpha - \frac{\varepsilon^2}{d} \int_0^1 \tilde{\theta}_{\alpha\alpha}^2 d\alpha + \frac{\varepsilon |b|}{d} \int_0^1 |\tilde{\theta}_a| |\tilde{\theta}_{\alpha\alpha}| d\alpha. \tag{7.35}
\]
Now, apply the Cauchy-Schwarz and Young's inequalities with \(\mu > 0\) and collect terms to find
\[
\dot{W} \leq -\left( d - \frac{|b|}{2m_0} \right) m_0 \tilde{\theta}^2(0) - \left( d - \frac{|b|}{2m_1} \right) m_1 \tilde{\theta}^2(1) \\
- \varepsilon \left( 1 - \frac{\mu |b|}{2d} \right) \int_0^1 \tilde{\theta}_a^2 d\alpha - \frac{\varepsilon^2}{d} \left( 1 - \frac{|b|}{2\mu \varepsilon} \right) \int_0^1 \tilde{\theta}_{\alpha\alpha}^2 d\alpha. \tag{7.36}
\]
The parameter \(\mu\) must satisfy the inequality,
\[
\frac{|b|}{2\varepsilon} < \mu < \frac{2d}{|b|}, \tag{7.37}
\]
so that the last two terms of (7.36) are negative. Since \(d > \frac{b^2}{4\varepsilon}\), such a choice of \(\mu\) is possible.

Splitting the first term on the second line of (7.36) and applying the Poincaré inequality yields
\[
\dot{W} \leq -\delta_0 m_0 \tilde{\theta}^2(0) - \delta_1 m_1 \tilde{\theta}^2(1) - \delta_2 \int_0^1 \tilde{\theta}^2 d\alpha \\
- \delta_3 \varepsilon \int_0^1 \tilde{\theta}_a^2 d\alpha - \frac{\varepsilon^2}{d} \left( 1 - \frac{|b|}{2\mu \varepsilon} \right) \int_0^1 \tilde{\theta}_{\alpha\alpha}^2 d\alpha, \\
\leq -2 \min (\delta_0, \delta_1, \delta_2, \delta_3) W, \tag{7.38}
\]
where
\[
\delta_0 = d - \frac{|b|}{2m_0}, \quad \delta_1 = d - \frac{1}{2m_1} \left( |b| + \frac{\varepsilon}{2} \left( 1 - \frac{\mu |b|}{2d} \right) \right), \\
\delta_2 = \frac{\varepsilon}{8} \left( 1 - \frac{\mu |b|}{2d} \right), \\
\delta_3 = \frac{d}{2} \left( 1 - \frac{\mu |b|}{2d} \right).
\]
We select the parameters,

\[ m_0 = \frac{|b|}{d}, \quad (7.39) \]

\[ m_1 = \frac{1}{d} \left( |b| + \frac{\varepsilon}{2} \left( 1 - \frac{\mu|b|}{2d} \right) \right), \quad (7.40) \]

to obtain \( \delta_0 = \delta_1 = \frac{d}{2} \). Noting, \( \delta_3 < \frac{d}{2} \), we have

\[ \dot{W} \leq -2 \min(\delta_2, \delta_3) W, \quad (7.41) \]

and by the Comparison Lemma [55],

\[ W(t) \leq e^{-2\eta t} W(0), \quad (7.42) \]

where \( \eta = \min(\delta_2, \delta_3) \). Finally, let \( p_0 = \min(m_0, m_1, 1, \varepsilon/d) \) and \( p_1 = \max(m_0, m_1, 1, \varepsilon/d) \). Then,

\[ \Gamma(t) \leq \frac{2}{p_0} W(t) \leq \frac{2}{p_0} e^{-2\eta t} W(0) \leq \frac{p_1}{p_0} e^{-2\eta t} \Gamma(0), \quad (7.43) \]

and the result (7.24) is obtained with \( M = p_1/p_0 \).

We now state a bound for \( \tilde{\theta}(\alpha) \).

**Lemma 7.3** \( \sup_{\alpha \in [0,1]} |\tilde{\theta}(\alpha)| \leq \sqrt{2\Gamma(t)} \)

**Proof:** The proof follows after noting \( \int_0^\alpha \partial_\alpha \left( \frac{1}{2} \tilde{\theta}^2 \right) d\alpha = \int_0^\alpha \tilde{\theta} \partial_\alpha \tilde{\theta} d\alpha \).

Stability of the \( r^2 \)-subsystem (7.13), (7.15) via Lyapunov analysis is difficult because the subsystem has two equilibria: 1) \( r^2 = 0 \), which is unstable, and 2) \( r^2 = R^2 \), which is stable. We could remove the equilibrium at the origin by modifying the control laws (7.5), (7.6), but in doing so, the control laws become singular at the origin. Instead, we consider agents in the domain,

\[ \mathcal{D} = \{(u, v) \in \mathbb{R}^2 \mid u^2 + v^2 \geq \rho^2, \rho > 0\}, \]

and utilize the explicit solution (7.17) to give the following point-wise convergence result, whose bounds depend on where the agent is initialized:

**Theorem 7.1** Each agent \( \alpha \) in the system (7.1)–(7.6), with boundary conditions (7.7), (7.8), exponentially converges to its equilibrium position on the circle
\( (x(\alpha,t), y(\alpha,t)) \equiv (R \cos \tilde{\theta}(\alpha), R \sin \tilde{\theta}(\alpha)) \) where \( R > 0 \) and \( \tilde{\theta}(\alpha) \) is given by (7.22) if \( b = 0 \) or by (7.23) if \( b \neq 0 \). Specifically, for all initial profiles such that \( x^2(\alpha,0) + y^2(\alpha,0) \geq \rho^2 > 0 \) the following bound holds for each agent:

\[
\max (|\bar{x}(\alpha,t)|, |\bar{y}(\alpha,t)|) \leq F(\alpha) e^{-\eta t} + G(\alpha) e^{-\eta t}, \quad \forall t \geq 0, \quad (7.44)
\]

where the deployment error variables are \( \bar{x}(\alpha,t) = x(\alpha,t) - R \cos \tilde{\theta}(\alpha) \) and \( \bar{y}(\alpha,t) = y(\alpha,t) - R \sin \tilde{\theta}(\alpha) \), and

\[
F(\alpha) = \begin{cases} \frac{R}{2 \rho^2} |r_0^2(\alpha) - R^2|, & r_0(\alpha) < R, \\ |r_0(\alpha) - R|, & r_0(\alpha) > R, \end{cases} \quad (7.45)
\]

\[
G(\alpha) = \sqrt{2MT(0)} R \left( |\cos \tilde{\theta}(\alpha)| + |\sin \tilde{\theta}(\alpha)| \right), \quad (7.46)
\]

\( M > 0 \), and \( \Gamma \) is given by (7.25).

**Proof:** We show the details of bounding \( \bar{x}(\alpha,t) \) only. We rewrite \( \bar{x}(\alpha,t) \) using (7.9), (7.17), (7.26), and some algebra as,

\[
\bar{x} = \frac{r_0 R}{\sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \cos \left( \tilde{\theta} + \tilde{\theta} \right) - R \cos \left( \tilde{\theta} + \tilde{\theta} \right) - R \cos \tilde{\theta} - R \cos \left( \tilde{\theta} + \tilde{\theta} \right)  \\
= \frac{r_0^2 R e^{-\eta R^2 t}}{\sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \left( \frac{r_0 + \sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}}{r_0 + \sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \right) \left( \cos \tilde{\theta} \cos \tilde{\theta} - \sin \tilde{\theta} \sin \tilde{\theta} \right)  \\
- R \cos \tilde{\theta} \left( 1 - \cos \tilde{\theta} \right) - R \sin \tilde{\theta} \sin \tilde{\theta}. \quad (7.47)
\]

From this expression, we bound \( |\bar{x}| \) using Young’s inequality and by noting \( |1 - \cos \tilde{\theta}| \leq |\tilde{\theta}| \) and \( |\sin \tilde{\theta}| \leq |\tilde{\theta}| \), which yields

\[
|\bar{x}| \leq \left| \frac{1}{\sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \right| \left| \frac{1}{r_0 + \sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \right| \left| \frac{r_0^2}{r_0^2} \right|  \\
\times \left( |\cos \tilde{\theta}| |\cos \tilde{\theta}| + |\sin \tilde{\theta}| |\sin \tilde{\theta}| \right) R e^{-\eta R^2 t}  \\
+ R |\cos \tilde{\theta}| |1 - \cos \tilde{\theta}| + R |\sin \tilde{\theta}| |\sin \tilde{\theta}|,  \\
= \left| \frac{1}{\sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \right| \left| \frac{1}{r_0 + \sqrt{r_0^2 - r_0^2 e^{-\eta R^2 t}}} \right| \left| \frac{r_0^2}{r_0^2} \right| R e^{-\eta R^2 t}  \\
+ R \left( |\cos \tilde{\theta}| + |\sin \tilde{\theta}| \right) |\tilde{\theta}|. \quad (7.48)
\]
We now proceed by bounding the first two terms of (7.48). The bounds depend on whether $\tilde{r}_0^2(\alpha)$ is positive or negative. When $r_0(\alpha) < R$, the first and second terms of (7.48) are bounded by

$$\left| \frac{1}{\sqrt{r_0^2 + (R^2 - r_0^2)} e^{-cR_0^2 t}} \right| \leq \frac{1}{r_0} \leq \frac{1}{\rho}, \quad (7.49)$$

$$\left| \frac{1}{r_0 + \sqrt{r_0^2 + (R^2 - r_0^2)} e^{-cR_0^2 t}} \right| \leq \frac{1}{2r_0} \leq \frac{1}{2\rho}. \quad (7.50)$$

If $r_0(\alpha) > R$, we have

$$\left| \frac{1}{\sqrt{r_0^2 - (R_0^2 - R^2)} e^{-cR_0^2 t}} \right| \leq \frac{1}{R}, \quad (7.51)$$

$$\left| \frac{1}{r_0 + \sqrt{r_0^2 - (R_0^2 - R^2)} e^{-cR_0^2 t}} \right| \leq \frac{1}{r_0 + R}. \quad (7.52)$$

Using inequalities (7.49)–(7.52) with Lemmas 7.2 and 7.3, we obtain

$$|\ddot{x}| \leq \frac{R}{2\rho^2} r_0^2 - R^2 \left| e^{-cR_0^2 t} + G(\alpha) e^{-\eta t} \right. \quad (7.53)$$

when $r_0(\alpha) < R$, and

$$|\ddot{x}| \leq |r_0 - R| e^{-cR_0^2 t} + G(\alpha) e^{-\eta t} \quad (7.54)$$

when $r_0(\alpha) > R$.

### 7.2 Deployment to Elliptical Arcs and Spirals

In the system (7.1)–(7.8), the anchor agents can manipulate only the follower agents’ $\bar{\theta}(\alpha)$-deployment family. The $r^2$-subsystem is an $\alpha$-parameterized family of ODEs, which means the radial deployment is decentralized and cannot be influenced by the anchor agents. By replacing the parameter $R$ in (7.5)–(7.8) with the variable $R(\alpha, t)$, which represents the desired radial position of agent $\alpha$ at time $t$, we provide the anchors the ability to manipulate the radial positions of the follower agents.
Consider the $R$-subsystem,

\[ R_t(\alpha, t) = \gamma R_{aa}(\alpha, t), \quad (7.55) \]
\[ R(0, t) = \bar{r}_0, \quad (7.56) \]
\[ R(1, t) = \bar{r}_1, \quad (7.57) \]

where $\gamma > 0$ and $\bar{r}_0$, $\bar{r}_1$ represent the desired radii for the anchor agents. This subsystem has a linear equilibrium profile that represents the agents’ desired radial positions, namely,

\[ \bar{R}(\alpha) = \bar{r}_0 + (\bar{r}_1 - \bar{r}_0) \alpha, \quad (7.58) \]

which one can show is exponentially stable in the $L^2$ norm. By setting $\bar{r}_0$ and $\bar{r}_1$, the anchor agents influence the radial deployment of the follower agents since $r^2(\alpha, t) \to R^2(\alpha, t) \to \bar{R}^2(\alpha)$ as $t \to \infty$. For consistency with $\bar{R}(\alpha)$, we require the anchor agents’ desired positions to satisfy $\bar{x}^2(l) + \bar{y}^2(l) = \bar{r}_l^2$.

The follower agents implement (7.55) to update the variable $R(\alpha, t)$, which replaces $R$ in the control laws (7.5), (7.6). Together, (7.5), (7.6), and (7.55) represent a dynamic feedback law for the follower agents. Thus, the user can change the radius of the formation by moving the anchors to new radial positions, deploy the agents to elliptic arcs by setting $\bar{r}_0 \neq \bar{r}_1$, and create spirals by encircling the origin with an anchor while increasing/decreasing its desired radius.

### 7.3 Discretized Agent Control Laws

Implementable control laws for a finite number of agents are obtained by spatially discretizing the control laws (7.5), (7.6) and if employed, the subsystem (7.55). When we spatially discretize the system, the state variables, $x(\alpha, t), y(\alpha, t)$, and $R(\alpha, t)$, become $x(ih, t), y(ih, t)$, and $R(ih, t)$, where $i = 0, \ldots, n$, $h = 1/n$, and $(n + 1)$ is the total number of agents. We denote the position and desired radius of the follower agents as $(x_i, y_i)$ and $R_i$, and the position and desired radius of the anchor agents as $(x_0, y_0)$, $R_0$ and $(x_n, y_n)$, $R_n$ respectively.

Using three-point central differencing to approximate the spatial derivatives in (7.5), (7.6), and (7.55), we obtain the discretized dynamic controller for follower
agent $i$,

$$
\hat{R}_i(t) = \frac{\gamma}{h^2} (R_{i+1} - 2R_i + R_{i-1}),
$$

(7.59)

$$
U_i(t) = -\frac{c}{2} x_i (x_i^2 + y_i^2 - R_i^2) - \varepsilon \frac{x_i y_i}{h^2} \frac{x_i + y_i - 2y_i + y_{i-1}}{x_i^2 + y_i^2} + \varepsilon \frac{y_i^2}{h^2} \frac{x_{i+1} - 2x_i + x_{i-1}}{x_i^2 + y_i^2} \\
- \varepsilon \frac{x_i y_i^2}{2h^2} \frac{(x_{i+1} - x_i - 1)^2}{(x_i^2 + y_i^2)^2} - (y_{i+1} - y_{i-1})^2 \\
+ \varepsilon \frac{x_i y_i}{2h^2} \frac{(x_{i+1} - x_i - 1)(y_{i+1} - y_{i-1})(x_i^2 - y_i^2)}{(x_i^2 + y_i^2)^2} \\
- \frac{b}{2h} y_i x_i (y_{i+1} - y_{i-1} - y_i (x_{i+1} - x_{i-1})),
$$

(7.60)

$$
V_i(t) = -\frac{c}{2} y_i (x_i^2 + y_i^2 - R_i^2) + \varepsilon \frac{x_i^2 y_i + 1 - 2y_i + y_{i-1}}{x_i^2 + y_i^2} - \varepsilon \frac{x_i y_i}{h^2} \frac{x_{i+1} - 2x_i + x_{i-1}}{x_i^2 + y_i^2} \\
+ \varepsilon \frac{x_i^2 y_i}{2h^2} \frac{(x_{i+1} - x_i - 1)^2}{(x_i^2 + y_i^2)^2} - (y_{i+1} - y_{i-1})^2 \\
- \varepsilon \frac{x_i}{2h^2} \frac{x_i (x_{i+1} - x_i - 1)(y_{i+1} - y_{i-1})(x_i^2 - y_i^2)}{(x_i^2 + y_i^2)^2} \\
+ \frac{b}{2h} x_i (y_{i+1} - y_{i-1} - y_i (x_{i+1} - x_{i-1})).
$$

(7.61)

If the agents stabilize only circular arcs, then we simply replace $R_i$ with the parameter $R$.

Due to the approximation error inherent in the spatial discretization, the achieved deployment will differ from the desired formations. The accuracy of the three-point central difference approximation is $O(h^2)$ [84], so as more agents are deployed, the approximation error will tend to zero. One should also note that the spatial discretization imposes a communication topology on the agents where each agent requires only local information.

For implementation purposes, we replace the term $\left(\arctan \frac{y_i}{x_i} - \arctan \frac{\bar{y} \bar{x}}{x_i}\right)$ in the anchor agents’ control laws (7.7), (7.8), with $\left(\arctan \frac{\bar{y} \bar{x}}{x_i} - \arctan \frac{\bar{y} \bar{x}}{x_i} - 2k^*\pi\right)$ where

$$
k^* = \arg \min_{k \in \{1,0,-1\}} \left| \arctan \frac{y_i}{x_i} - \arctan \frac{\bar{y} \bar{x}}{x_i} - 2k\pi \right|.
$$

(7.62)

This modification is best explained by example. Let $\theta(0,t) = \pi + \sigma$ and $\bar{\theta}_0 = \pi - \sigma$ where $\sigma$ is small. Without the modification, the anchor’s control law in polar
coordinates would be $\theta_t(0, t) = -2d(\sigma - \pi)$, causing it to move counterclockwise about the origin, when $\theta_t(0, t) = -2d\sigma$ is the desired response, which is achieved with $k^* = 1$. This behavior is not a concern with the follower agents since there are no trigonometric terms in their feedback laws.

### 7.4 Simulations

We present numerical examples that highlight how the anchor agents can be used to maneuver the followers from one formation to another defined by the nonlinear PDE model. We deploy 15 agents with gains $c = 0.3$, $\varepsilon = 2$, $b = -5$, $d = 5$, and $\gamma = 1$. Note, the negative value of $b$ means there will be a higher density of agents near the anchor at $\alpha = 0$. In the figures, darker shading represents an advance in time, and the total time elapsed is 2 seconds. A thin solid line represents the agents’ initial positions, and the final positions are shaded green.

The agents first stabilize a circular arc from $\theta_0 = 0$ rad to $\theta_1 = \pi/2$ rad with the anchor agents setting $\bar{r}_0 = \bar{r}_n = 5$ m (Figure 7.2). The variable $R(\alpha, t)$ is initialized from a uniform distribution sampled from the interval $[2, 8]$. From this deployment, the anchor agents change parameters to $\theta_0 = \theta_1 = -2\pi/3$ rad and $\bar{r}_0 = \bar{r}_n = 8$ m, which causes them to transverse the circle and bring the follower agents into a circle deployment (Figure 7.3(a)) with nonuniform coverage due to the advection term in (7.14). Finally, we set $\theta_0 = \pi/2$ rad, $\theta_1 = -\pi/4$ rad,
\( \bar{r}_0 = 15 \text{ m}, \) and \( \bar{r}_n = 5 \text{ m} \) to move the agents from the circle formation to a spiral deployment (Figure 7.3(b)).

7.5 Summary

We have presented a nonlinear PDE model (7.1)–(7.6) that enables follower agents to deploy to arc formations that can be rotated, stretched, or shortened by moving anchor agents about the arc with controllers (7.7), (7.8). A circle formation is formed by collocating the anchors but is not guaranteed since the system has multiple equilibria. If the initial condition dictates that the agents, instead, rendezvous at the anchors’ position, the anchors can transverse the circle in opposite directions until they are again collocated, dragging the followers into a circle formation.

By replacing the parameter \( R \) with the \( R \)-subsystem (7.55)–(7.57), the follower agents are no longer limited to formations of fixed radius and can deploy to formations with the desired radial profile (7.58). Thus, a user can deploy the followers, without changing their controllers, to formations of practical interest—arcs, circles, and spirals—by manipulating the formation geometry with the anchor...
agents.

This chapter is an adaptation of material appearing in: P. Frihauf and M. Krstic, “Multi-agent deployment to a family of planar arcs,” in *Proc. of American Control Conference*, Baltimore, MD, 2010.

The dissertation author is the primary investigator and author of this work.
Appendix A

Integrals for Computing Average Error Systems

For the convergence analysis of Nash seeking strategies, average systems for \(N\)-player and infinitely-many player games are computed. For \(N\)-player games, integrals over the period \(T = 2\pi \times \text{LCM}\{1/\bar{\omega}_1, \ldots, 1/\bar{\omega}_N\}\), where LCM denotes the least common multiple, are evaluated, and for games with infinitely-many players, integrals are evaluated according to the general averaging theory [45].

Section A.1 details the necessary calculations to compute the average error system for games with quadratic payoff functions, Section A.2 presents the additional calculations needed for games with non-quadratic payoff functions, and Section A.3 shows the calculations needed to compute the average error system for infinitely-many player games. To obtain the average systems used in the analysis of Newton-based Nash seeking scheme in Chapter ??, integrals of the forms shown in Sections A.1 and A.2 are computed.

A.1 Average Error System for Games with Quadratic Payoffs

To obtain (2.22), we compute the average of each term in the error system (2.21) and assume that \(\omega_i \neq \omega_j\), \(2\omega_i \neq \omega_j\) and \(\omega_i \neq \omega_j + \omega_k\) for all distinct
\[ i, j, k \in \{1, \ldots, N\}. \] The averages of the first and last terms of (2.21) are zero since
\[
\frac{1}{T} \int_0^T \mu_i(\tau) \, d\tau = a_i \frac{\cos \varphi_i - \cos(\bar{\omega}_i T + \varphi_i)}{\bar{\omega}_i} = 0. \tag{A.1}
\]
The average of the second term of (2.21) depends on whether or not \(i = k\). If \(i \neq k\),
\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu_i(\tau) \mu_k(\tau) \, d\tau = a_i a_k \lim_{T \to \infty} \frac{1}{T} \int_0^T \sin(\gamma_i \tau + \varphi_i) \sin(\gamma_k \tau + \varphi_k) \, d\tau,
\]
\[
= a_i a_k \lim_{T \to \infty} \frac{1}{2T} \int_0^T \left[ \cos((\gamma_i - \gamma_k) \tau + \varphi_i - \varphi_k) - \cos((\gamma_i + \gamma_k) \tau + \varphi_i + \varphi_k) \right] \, d\tau,
\]
\[
= 0, \tag{A.2}
\]
since \(\omega_i \neq \omega_k\), and if \(i = k\),
\[
\frac{1}{T} \int_0^T \mu_i^2(\tau) \, d\tau = \frac{a_i^2}{T} \int_0^T \sin^2(\bar{\omega}_i \tau + \varphi_i) \, d\tau,
\]
\[
= \frac{a_i^2}{2T} \int_0^T [1 - \cos(2\bar{\omega}_i \tau + 2\varphi_i)] \, d\tau
\]
\[
= \frac{a_i^2}{2}, \tag{A.3}
\]
which with (A.2) implies
\[
\frac{1}{T} \int_0^T \frac{k_i}{\omega} \sum_{j=1}^N \sum_{k=1}^N \delta_{jk} (\bar{u}_j(\tau) + u_j^*) \mu_i(\tau) \mu_k(\tau) \, d\tau = \frac{k_i a_i^2}{2\omega} \sum_{j=1}^N \delta_{ij} (\bar{u}_j(\tau) + u_j^*). \tag{A.4}
\]
Similarly, the average of the fourth term of (2.21) is
\[
\frac{1}{T} \int_0^T \frac{k_i}{\omega} \sum_{j=1}^N \delta_{j} (\bar{u}_j(\tau) + u_j^* + \mu_j(\tau)) \mu_i(\tau) \, d\tau = \frac{k_i a_i^2}{2\omega} d_i. \tag{A.5}
\]
The average of the third term requires computing the following three integrals:

$$\frac{1}{T} \int_0^T \mu^3_i(\tau) d\tau$$

$$= \frac{a_i^3}{T} \int_0^T \sin^3(\bar{\omega}_i \tau + \varphi_i) d\tau,$$

$$= \frac{3a_i^3}{4T} \int_0^T \sin(\bar{\omega}_i \tau + \varphi_i) d\tau - \frac{a_i^3}{4T} \int_0^T \sin(3\bar{\omega}_i \tau + 3\varphi_i)) d\tau,$$

$$= 0,$$  \hfill (A.6)

$$\frac{1}{T} \int_0^T \mu^2_i(\tau) \mu_j(\tau) d\tau$$

$$= \frac{a_i^2 a_j}{T} \int_0^T \sin^2(\bar{\omega}_i \tau + \varphi_i) \sin(\bar{\omega}_j \tau + \varphi_j) d\tau,$$

$$= \frac{a_i^2 a_j}{2T} \int_0^T [\sin(\bar{\omega}_j \tau + \varphi_j) - \cos(2\bar{\omega}_i \tau + 2\varphi_i) \sin(\bar{\omega}_j \tau + \varphi_j)] d\tau,$$

$$= \frac{a_i^2 a_j}{4T} \int_0^T [2 \sin(\bar{\omega}_j \tau + \varphi_j) - \sin((2\bar{\omega}_i + \bar{\omega}_j) \tau + 2\varphi_i + \varphi_j)$$

$$+ \sin((\bar{\omega}_i - \bar{\omega}_j) \tau + 2\varphi_i - \varphi_j)] d\tau,$$

$$= 0,$$  \hfill (A.7)

$$\frac{1}{T} \int_0^T \mu_i(\tau) \mu_j(\tau) \mu_k(\tau) d\tau$$

$$= \frac{a_i a_j a_k}{T} \int_0^T \sin(\bar{\omega}_i \tau + \varphi_i) \sin(\bar{\omega}_j \tau + \varphi_j) \sin(\bar{\omega}_k \tau + \varphi_k) d\tau,$$

$$= \frac{a_i a_j a_k}{4T} \int_0^T [\sin((\bar{\omega}_i - \bar{\omega}_j + \bar{\omega}_k) \tau + \varphi_i - \varphi_j + \varphi_k)$$

$$- \sin((\bar{\omega}_i - \bar{\omega}_j - \bar{\omega}_k) \tau + \varphi_i - \varphi_j - \varphi_k)$$

$$- \sin((\bar{\omega}_i + \bar{\omega}_j + \bar{\omega}_k) \tau + \varphi_i + \varphi_j + \varphi_k)$$

$$+ \sin((\bar{\omega}_i + \bar{\omega}_j - \bar{\omega}_k) \tau + \varphi_i + \varphi_j - \varphi_k)] d\tau,$$

$$= 0,$$  \hfill (A.8)

where we have assumed $2\omega_i \neq \omega_j$ and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$. 
A.2 Average Error System for Games with Non-Quadratic Payoffs

For games with non-quadratic payoffs, we assume the players’ frequencies satisfy the requirements for games with quadratic payoff functions and also \(3\omega_i \neq \omega_j, \omega_i \neq 2\omega_j + \omega_k, 2\omega_i \neq \omega_j + \omega_k\), where \(i, j, k \in \{1, \ldots, N\}\) are distinct. Then, the following integrals, in addition to (A.1), (A.2), (A.3), (A.6), (A.7), and (A.8), are computed to obtain (3.16):

\[
\frac{1}{T} \int_0^T \mu_i^4(\tau) d\tau = \frac{a_i^4}{8T} \int_0^T [3 - 4\cos(2\bar{\omega}_i\tau + 2\varphi_i) + \cos(4\bar{\omega}_i\tau + 4\varphi_i)] d\tau, \\
= \frac{3a_i^4}{8}, \quad \text{(A.9)}
\]

\[
\frac{1}{T} \int_0^T \mu_i^3(\tau)\mu_j(\tau) d\tau = \frac{a_i^3a_j}{4T} \int_0^T \left[ 3\sin(\bar{\omega}_i\tau + \varphi_j)\sin(\bar{\omega}_j\tau + \varphi_j) \\
- \sin(3\bar{\omega}_i\tau + 3\varphi_i)\sin(\bar{\omega}_j\tau + \varphi_j) \right] d\tau, \\
= \frac{a_i^3a_j}{8T} \int_0^T \left[ 3\cos((\bar{\omega}_i - \bar{\omega}_j)\tau + \varphi_i - \varphi_j) \\
- 3\cos((\bar{\omega}_i + \bar{\omega}_j)\tau + \varphi_i + \varphi_j) \\
- \cos((3\bar{\omega}_i - \bar{\omega}_j)\tau + 3\varphi_i - \varphi_j) \\
+ \cos((3\bar{\omega}_i + \bar{\omega}_j)\tau + 3\varphi_i + \varphi_j) \right] d\tau, \\
= 0, \quad \text{(A.10)}
\]

\[
\frac{1}{T} \int_0^T \mu_i^2(\tau)\mu_j^2(\tau) d\tau = \frac{a_i^2a_j^2}{8T} \int_0^T \left[ 2 - 2\cos(2\bar{\omega}_i\tau + 2\varphi_i) - 2\cos(2\bar{\omega}_j\tau + 2\varphi_j) \\
+ \cos(2(\bar{\omega}_i - \bar{\omega}_j)\tau + 2(\varphi_i - \varphi_j)) \\
+ \cos(2(\bar{\omega}_i + \bar{\omega}_j)\tau + 2(\varphi_i + \varphi_j)) \right] d\tau, \\
= \frac{a_i^2a_j^2}{4}, \quad \text{(A.11)}
\]

\[
\frac{1}{T} \int_0^T \mu_i(\tau)\mu_j^2(\tau)\mu_k(\tau) d\tau = \frac{a_i a_j^2 a_k}{2T} \int_0^T \sin(\bar{\omega}_i\tau + \varphi_i) \\
\times (1 - \cos(2\bar{\omega}_j\tau + 2\varphi_j))\sin(\bar{\omega}_k\tau + \varphi_k) d\tau, \\
= \frac{a_i a_j^2 a_k}{4T} \int_0^T \cos((\bar{\omega}_i - \bar{\omega}_k)\tau + \varphi_i - \varphi_k)
\]
\[-\cos((\bar{\omega}_i + \bar{\omega}_k)\tau + \varphi_i + \varphi_k)] \\
\times (1 - \cos(2\bar{\omega}_j\tau + 2\varphi_j))
\]
\[= \frac{a_i a_j a_k}{8T} \int_0^T [2\cos((\bar{\omega}_i - \bar{\omega}_k)\tau + \varphi_i - \varphi_k) \\
- 2\cos((\bar{\omega}_i + \bar{\omega}_k)\tau + \varphi_i + \varphi_k) \\
- \cos((\bar{\omega}_i - 2\bar{\omega}_j - \bar{\omega}_k)\tau + \varphi_i - 2\varphi_j - \varphi_k) \\
- \cos((\bar{\omega}_i + 2\bar{\omega}_j - \bar{\omega}_k)\tau + \varphi_i + 2\varphi_j - \varphi_k) \\
+ \cos((\bar{\omega}_i - 2\bar{\omega}_j + \bar{\omega}_k)\tau + \varphi_i - 2\varphi_j + \varphi_k) \\
+ \cos((\bar{\omega}_i + 2\bar{\omega}_j + \bar{\omega}_k)\tau + \varphi_i + 2\varphi_j + \varphi_k)] d\tau,
\]
\[= 0. \quad (A.12)\]

\section{A.3 Average Error System for Games with Infinitely-Many Players}

To obtain the average error system (4.15), we compute the average of the operator \(G \ (4.14)\), which requires averaging terms of the following forms: \(\mu(x, t)\), \(\mu^2(x, t)\), \(\mu^3(x, t)\), \(\mu(x, t) \int_0^1 \rho(y) \mu(y, t) dy\), and \(\mu^2(x, t) \int_0^1 \rho(y) \mu(y, t) dy\), where \(\rho\) is bounded and measurable.

Computing the average of the first three terms is straightforward:

\[\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu(x, t) dt = a(x) \lim_{T \to \infty} \frac{\cos \varphi(x) - \cos(\omega(x)T + \varphi(x))}{T \omega(x)},\]
\[= 0, \quad (A.13)\]

\[\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu^2(x, t) dt = \lim_{T \to \infty} \frac{a^2(x)}{2T} \int_0^T [1 - \cos(2\omega(x)t + 2\varphi(x))] dt,
\]
\[= \frac{a^2(x)}{2}, \quad (A.14)\]

\[\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu^3(x, t) dt = \lim_{T \to \infty} \frac{3a^3(x)}{4T} \int_0^T \sin(\omega(x)t + \varphi(x)) dt
\]
\[\quad - \lim_{T \to \infty} \frac{a^3(x)}{4T} \int_0^T \sin(3\omega(x)t + 3\varphi(x)) dt,
\]
\[= 0. \quad (A.15)\]
Calculating the averages of $\mu(x, t) \int_0^1 \rho(y) \mu(y, t) dy$ and $\mu^2(x, t) \int_0^1 \rho(y) \mu(y, t) dy$ requires more care. We start by computing the average of $\mu(x, t) \int_0^1 \rho(y) \mu(y, t) dy$,

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu(x, t) \int_0^1 \rho(y) \mu(y, t) dy \; dt
= a(x) \lim_{T \to \infty} \int_0^1 \frac{a(y) \rho(y)}{T} \int_0^T \sin(\omega(x) t + \varphi(x)) \sin(\omega(y) \tau + \varphi(y)) \; d\tau \; dy,
$$

(A.16)

where we have switched the order of integration and substituted for $\mu(x, t)$ and $\mu(y, t)$. Next, we compute the inner integral over $t$ on the last line of (A.16). Specifically,

$$
\frac{1}{T} \int_0^T \sin(\omega(x) t + \varphi(x)) \sin(\omega(y) t + \varphi(y)) \; dt
= \frac{1}{2T} \int_0^T \left[ \cos((\omega(x) - \omega(y)) t + \varphi(x) - \varphi(y))
- \cos((\omega(x) + \omega(y)) t + \varphi(x) + \varphi(y)) \right] \; dt,
$$

$$
= \frac{1}{2} \left( \frac{\sin((\omega(x) - \omega(y)) T + \varphi(x) - \varphi(y))}{(\omega(x) - \omega(y)) T}
- \frac{\sin(\varphi(x) - \varphi(y))}{(\omega(x) - \omega(y)) T}
+ \frac{\sin(\varphi(x) + \varphi(y))}{(\omega(x) + \omega(y)) T}
- \frac{\sin((\omega(x) + \omega(y)) T + \varphi(x) + \varphi(y))}{(\omega(x) + \omega(y)) T} \right),
$$

(A.17)

where we have used the fact that for any given $x$, the set $\{ y \in [0, 1] \mid \omega(x) = \omega(y) \}$ is of measure zero since $\omega \in \Omega$.

To switch the order of the limit and the integration over $y$ in (A.16), we apply the dominated convergence theorem, which requires that the integrand $a(y) \rho(y) \delta_0(x, y, T)$ be bounded by a function $\eta_0(x, y)$ and that $\int_0^1 \eta_0(x, y) dy$ be finite. Using the sum of angles trigonometric identity, we have

$$
|\delta_0(x, y, T)| \leq \frac{1}{2} \left( |\cos(\varphi(x) - \varphi(y))| \left| \frac{\sin((\omega(x) - \omega(y)) T)}{(\omega(x) - \omega(y)) T} \right|
+ |\sin(\varphi(x) - \varphi(y))| \left| \frac{\cos((\omega(x) - \omega(y)) T) - 1}{(\omega(x) - \omega(y)) T} \right|
+ |\cos(\varphi(x) + \varphi(y))| \left| \frac{\sin((\omega(x) + \omega(y)) T)}{(\omega(x) + \omega(y)) T} \right| \right)
$$
\[+ \left| \sin(\varphi(x) + \varphi(y))\right| \left| \frac{\cos((\omega(x) + \omega(y))T) - 1}{(\omega(x) + \omega(y))T} \right|,\]

\[\leq 2, \tag{A.18}\]

which implies the following bound on the integrand,

\[|a(y)\rho(y)\delta_0(x, y, T)| \leq 2 \max_{y \in [0, 1]} \{a(y)\rho(y)\}. \tag{A.19}\]

Clearly, \(\int_0^1 2 \max_{y \in [0, 1]} \{a(y)\rho(y)\} dy < \infty\), which with (A.19), allows the dominated convergence theorem to be applied to (A.16). Thus,

\[\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu(x, t) \int_0^1 \rho(y)\mu(y, t)dy\, dt = a(x) \int_0^1 \lim_{T \to \infty} a(y)\rho(y)\delta_0(x, y, T)dy,\]

\[= 0, \tag{A.20}\]

since \(\lim_{T \to \infty} \delta_0(x, y, T) = 0\).

The average of \(\mu^2(x, t)\int_0^1 \rho(y)\mu(y, t)dy\) is computed in a similar manner. We have

\[\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu^2(x, t) \int_0^1 \rho(y)\mu(y, t)dy\, dt = a^2(x) \int_0^1 \frac{a(y)\rho(y)}{T} \int_0^T \sin^2(\omega(x)t + \varphi(x)) \sin(\omega(y)t + \varphi(y))\, dt\, dy, \tag{A.21}\]

and computing the inner integral on the last line of (A.21) gives

\[\frac{1}{T} \int_0^T \sin^2(\omega(x)t + \varphi(x)) \sin(\omega(y)t + \varphi(y))\, dt\]

\[= \frac{1}{2T} \int_0^T [\sin(\omega(y)t + \varphi(y)) - \cos(2\omega(x)t + 2\varphi(x)) \sin(\omega(y)t + \varphi(y))]\, dt,\]

\[= \frac{1}{4T} \int_0^T [2\sin(\omega(y)t + \varphi(y)) - \sin((2\omega(x) + \omega(y))t + 2\varphi(x) + \varphi(y)) + \sin((2\omega(x) - \omega(y))t + 2\varphi(x) - \varphi(y))]\, dt,\]

\[= \frac{1}{4} \left( \frac{2 \cos(\varphi(y)) - 2 \cos(\omega(y)T + \varphi(y))}{\omega(y)T} \right)\]

\[+ \frac{\cos(2\omega(x) + \omega(y)T + 2\varphi(x) + \varphi(y))}{(2\omega(x) + \omega(y))T} - \frac{\cos(2\varphi(x) + \varphi(y))}{(2\omega(x) + \omega(y))T}\]

\[+ \frac{\cos(2\varphi(x) - \varphi(y))}{(2\omega(x) - \omega(y))T} - \frac{\cos(2\omega(x) - \omega(y)T + 2\varphi(x) - \varphi(y))}{(2\omega(x) - \omega(y))T}\),

\[= \delta_1(x, y, T), \tag{A.22}\]
where we have used the fact that for any given \( x \), the set \( \{ y \in [0, 1] \mid 2\omega(x) = \omega(y) \} \) is of measure zero since \( \omega \in \Omega \).

As before, to switch the order of the limit and the integration over \( y \) in (A.21), we apply the dominated convergence theorem. Bounding \( \delta_1(x, y, T) \) leads to

\[
|\delta_1(x, y, T)| \leq \frac{1}{4} \left( 2 |\cos\varphi(y)| \left| \frac{1 - \cos(\omega(y)T)}{\omega(y)T} \right| + 2 |\sin\varphi(y)| \left| \frac{\sin(\omega(y)T)}{\omega(y)T} \right| \right.
\]
\[
+ |\cos(2\varphi(x) + \varphi(y))| \left| \frac{\cos((2\omega(x) + \omega(y))T) - 1}{(2\omega(x) + \omega(y))T} \right| 
\]
\[
+ |\sin(2\varphi(x) + \varphi(y))| \left| \frac{\sin((2\omega(x) + \omega(y))T)}{(2\omega(x) + \omega(y))T} \right| 
\]
\[
+ |\cos(2\varphi(x) - \varphi(y))| \left| \frac{1 - \cos((2\omega(x) - \omega(y))T)}{(2\omega(x) - \omega(y))T} \right| 
\]
\[
+ |\sin(2\varphi(x) - \varphi(y))| \left| \frac{\sin((2\omega(x) - \omega(y))T)}{(2\omega(x) - \omega(y))T} \right| \),
\]

\[
\leq 2,
\]

which implies the bound on the integrand

\[
|a(y)\rho(y)\delta_1(x, y, T)| \leq 2 \max_{y \in [0,1]} a(y)\rho(y). \tag{A.24}
\]

Thus, the dominated convergence theorem applies and (A.21) becomes

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu^2(x, \tau) \int_0^1 \rho(y)\mu(y, t) \, dt = a^2(x) \int_0^1 \lim_{T \to \infty} a(y)\rho(y)\delta_1(x, y, T) \, dy,
\]

\[
= 0,
\]

\[
\tag{A.25}
\]

since \( \lim_{T \to \infty} \delta_1(x, y, T) = 0 \). From (A.13), (A.14), (A.15), (A.20), and (A.25), we obtain the average system (4.15).
Appendix B

Lemmas for Discrete-Time Extremum Seeking

The following lemmas are used to manipulate the error difference equation (5.27) into a form amenable for stability analysis.

Lemma B.1 For the transfer function $G(z)$ the following is true for any real $\varphi$:

$$G(z)[\cos(\omega k - \varphi)v(k)] = \text{Re}\{e^{j(\omega k - \varphi)}G(e^{j\omega}z)[v(k)]\}. \quad (B.1)$$

Lemma B.2 For any two rational functions $A(\cdot)$ and $B(\cdot, \cdot)$, the following is true:

$$\text{Re}\{e^{j(\omega_1k - \varphi_1)}A(e^{j\omega_1})\text{Re}\{e^{j(\omega_2k - \varphi_2)}B(z, e^{j\omega_2})[v(k)]\}$$

$$= \frac{1}{2}\text{Re}\{e^{j((\omega_2 - \omega_1)k + \varphi_1 - \varphi_2)}A(e^{-j\omega_1})B(z, e^{j\omega_2})[v(k)]\}$$

$$+ \frac{1}{2}\text{Re}\{e^{j((\omega_1 + \omega_2)k - \varphi_1 - \varphi_2)}A(e^{j\omega_1})B(z, e^{j\omega_2})[v(k)]\}. \quad (B.2)$$

Lemma B.3 For any rational function $B(\cdot, \cdot)$, the following is true:

$$\text{Re}\{e^{j(\omega k - \varphi)}B(z, e^{j\omega})[v(k)]\} = \cos(\omega k - \varphi)\text{Re}\{B(z, e^{j\omega})[v(k)]\}$$

$$- \sin(\omega k - \varphi)\text{Im}\{B(z, e^{j\omega})[v(k)]\}. \quad (B.3)$$
Appendix C

Lemmas for Leader-Enabled Deployment

Lemma C.1 The inequalities,

\[ p_1 \Omega(t) \leq \Psi(t) \leq p_2 \Omega(t), \]
\[ q_1 \Psi(t) \leq V(t) \leq q_2 \Psi(t), \]

hold, where \( \Omega(t) \) and \( V(t) \) are given by (6.45) and (6.46),

\[ \Psi(t) = |w(0, t)|^2 + |w(1, t)|^2 + \int_0^1 |w(\alpha, t)|^2 d\alpha + \int_0^1 |w_\alpha(\alpha, t)|^2 d\alpha, \]

and

\[ p_1 = \frac{1}{\max \left\{ 2 + \frac{|b|^2}{\varepsilon^2} + \left(12 + \frac{\varepsilon^2}{2}\right)L + 8L_\alpha, 4 \right\} \max \left\{ 1, \left| e^{-\frac{b}{2\varepsilon}} \right|^2 \right\}}, \] (C.1)
\[ p_2 = \max \left\{ 2 + \frac{|b|^2}{\varepsilon^2} + 8K + 4K_\alpha, 4 \right\} \max \left\{ 1, \left| e^{-\frac{b}{2\varepsilon}} \right|^2 \right\}, \] (C.2)
\[ q_1 = \frac{1}{16} \min \{1, 8m\}, \] (C.3)
\[ q_2 = \frac{1}{2} \max \{1, m\}, \] (C.4)

\[ L = \sup_{(\alpha, \beta) \in \mathcal{T}} |l(\alpha, \beta)|^2, \]
\[ L_\alpha = \sup_{(\alpha, \beta) \in \mathcal{T}} |l_\alpha(\alpha, \beta)|^2, \]
\[ K = \sup_{(\alpha, \beta) \in \mathcal{T}} |k(\alpha, \beta)|^2, \]
\[ K_\alpha = \sup_{(\alpha, \beta) \in \mathcal{T}} |k_\alpha(\alpha, \beta)|^2. \]
**Proof:** Using (6.27), (6.31), (6.36), and applying the Cauchy-Schwarz and Young’s inequalities several times to bound terms in \( \Psi(t) \) and \( \Omega(t) \), one can obtain (C.1) and (C.2). The scalar \( q_2 \) is immediate from \( V(t) \) and \( \Psi(t) \), and \( q_1 \) is found by splitting the last term of \( V(t) \) and applying the Poincaré inequality. ■

**Lemma C.2** The inequalities,

\[
\begin{align*}
    r_1 \Sigma(t) &\leq \Phi(t) \leq r_2 \Sigma(t), \\
    \Phi(t) &\leq W(t) \leq s_2 \Phi(t),
\end{align*}
\]

hold where \( \Sigma(t) \) and \( W(t) \) are given by (6.69) and (6.70),

\[
\Phi(t) = |\tilde{w}(0,t)|^2 + |\tilde{w}(1,t)|^2 + \int_0^1 |\tilde{w}(\alpha,t)|^2 \, d\alpha + \int_0^1 |\tilde{w}_\alpha(\alpha,t)|^2 \, d\alpha,
\]

and

\[
\begin{align*}
    r_1 &= \frac{1}{\max \left\{ 2 + \left| \frac{b}{\varepsilon} \right|^2 + \left( 12 + \left| \frac{b}{\varepsilon} \right|^2 \right) P + 8P_\alpha, \ 4 \right\} \max \left\{ 1, \left| e^{-\frac{b}{2\varepsilon}} \right|^2 \right\}}, \quad (C.5) \\
    r_2 &= \max \left\{ 2 + \left| \frac{b}{\varepsilon} \right|^2 + 8Q + 4Q_\alpha, \ 4 \right\} \max \left\{ 1, \left| e^{\frac{b}{2\varepsilon}} \right|^2 \right\}, \quad (C.6) \\
    s_1 &= \frac{1}{16}, \quad (C.7) \\
    s_2 &= \frac{1}{2^7}, \quad (C.8)
\end{align*}
\]

\[
P = \sup_{(\alpha,\beta) \in \mathcal{S}} |p(\alpha,\beta)|^2, \quad P_\alpha = \sup_{(\alpha,\beta) \in \mathcal{S}} |p_\alpha(\alpha,\beta)|^2, \quad Q = \sup_{(\alpha,\beta) \in \mathcal{S}} |q(\alpha,\beta)|^2, \quad Q_\alpha = \sup_{(\alpha,\beta) \in \mathcal{S}} |q_\alpha(\alpha,\beta)|^2.
\]

**Proof:** The proof is analogous to the proof of Lemma C.1. ■
Bibliography


