DISTRIBUTED TARGET TRACKING IN NONLINEAR MULTI-AGENT SYSTEMS

By

CRISTIAN F. NINO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2025

To my mother, father, and brother, thank you for your support, for instilling ambition, and for inspiring perseverance.

ACKNOWLEDGEMENTS

I am grateful to Dr. Warren E. Dixon for fostering academic freedom, intellectual rigour, and the habit of asking questions; to the U.S. Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship for supporting me throughout my M.S. and Ph.D. studies; and to the staff at Eglin Air Force Base (AFB), with whom I worked each summer, for their collaboration and support.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS 4

LIST OF TABLES 7

LIST OF FIGURES 8

ABSTRACT 9

CHAPTER 1 INTRODUCTION 10

1.1 Background 11
1.1.1 Challenges in Sensing and Communication 12
1.1.2 Control of Uncertain Nonlinear Systems 14
1.2 Dissertation Scope 16

CHAPTER 2 PRELIMINARIES 18

2.1 Notation and Mathematical Conventions 18
2.1.1 General and Set Notation 18
2.1.2 Vector and Matrix Notation 19
2.1.3 Function Spaces and Nonsmooth Analysis 20
2.1.4 Algebraic Graph Theory 21
2.1.5 Deep Residual Neural Network Model 22

CHAPTER 3 FOUNDATIONAL METHODS 25

3.1 Online ResNet-Based Adaptive Control for Nonlinear Target Tracking 25
3.1.1 Problem Formulation 25
3.1.1.1 Control Synthesis 27
3.1.1.2 Residual Neural Network Function Approximation 27
3.1.1.3 Control Design 31
3.1.2 Stability Analysis 31
3.1.3 Experiment 37
3.2 Second-Order Heterogeneous Multi-Agent Target Tracking without Relative Velocities 39
3.2.1 Problem Formulation 40
3.2.2 Control Synthesis 41
3.2.2.1 Distributed Observer Development 42
3.2.2.2 Distributed Controller Development 43
3.2.3 Stability Analysis 45
3.2.4 Simulation 48

CHAPTER 4 DISTRIBUTED RISE-BASED CONTROL FOR EXPONENTIAL HETEROGENEOUS MULTI-AGENT TARGET TRACKING OF SECOND-ORDER NONLINEAR SYSTEMS 55

4.1 Problem Formulation 55
4.1.1 System Dynamics 55
4.1.2 Control Objective 56
4.2 Control Design 57
4.3 Stability Analysis 61
4.4 Simulation. 67

CHAPTER 5 COLLABORATIVE STATE ESTIMATION AND TRACKING UNDER DIRECTED PARTIAL RELATIVE STATE FEEDBACK USING DEEP RESIDUAL NEURAL NETWORKS 70

5.1 Problem Formulation 70
5.1.1 System Dynamics 70
5.1.2 Communication Topology 72
5.1.3 Control Objective 72
5.2 Control Design 73
5.2.1 Residual Neural Network Function Approximation 74
5.2.2 Distributed Observer-based Control Design 75
5.2.3 Ensemble Representation 77
5.3 Stability Analysis 79
5.3.1 Trackability 80
5.3.2 Closed-Loop Error System. 86
5.4 Spacecraft Servicing Simulation 96
5.5 Conclusion 105

CHAPTER 6 CONCLUSIONS AND FUTURE WORK. 107

6.1 Conclusions 107
6.2 Future Work: Toward a Unified Framework for Multi-Agent Systems 107
6.2.1 Immediate Generalizations: Multi-Target and Formation Objectives. 108
6.2.2 A General Language: Cellular Sheaves for Heterogeneity. 108
6.2.3 Incorporating Spatiotemporal Dynamics 109

LIST OF REFERENCES. 110

BIOGRAPHICAL SKETCH 117

LIST OF TABLES

page

Table 3-1. Parameters used in the comparative experiment of SNN, DNN, and ResNet -based adaptive controllers. 37

Table 3-2. Result Comparison 53

Table 3-3. Heterogeneity Parameters for $\omega = 2.5$ 54

Table 5-1. Spacecraft Simulation Parameters 100

LIST OF FIGURES

page

Figure 2-1. Deep Residual Neural Network Architecture. 23

Figure 3-1. Tracking error comparison over the 360-second experiment. 38

Figure 3-2. Trajectory comparison over the 360-second experiment. 39

Figure 3-3. Block diagram of algorithm implementation for agent $i \in \mathcal{V}$. Control gains have been omitted for clarity. 45

Figure 3-4. Plots of the tracking and velocity estimation errors for the homogeneous comparison ($\omega = 0$). 51

Figure 3-5. Plot of the tracking errors for each agent, when the dynamic parameters a_i, b_i, c_i, d_i are selected from the uniform distribution $U(0, 2.5)$. 52

Figure 3-6. Comparative plots of the average tracking errors, when the dynamic parameters a_i, b_i, c_i, d_i are selected from the uniform distribution $U(0, 10)$. 52

Figure 4-1. Left: Tracking error norm for $t \in [0, 1.5]$. Right: Control effort norm for $t \in [0, 1.5]$s. 68

Figure 5-1. Communication topology visualization of the multi-spacecraft system. 98

Figure 5-2. 3-D visualization of the trajectories servicing agents (solid lines) and the defunct spacecraft (dotted line) during the first five seconds of the simulation. 99

Figure 5-3. Plot showing the norm of the velocity estimation error, $\left\|\tilde{\zeta}_i\right\|$, over time for all agents $i \in \mathcal{V}$. 101

Figure 5-4. Plot showing the norm of the tracking error, $\|e_i\|$, over time for all agents $i \in \mathcal{V}$. . . 103

Figure 5-5. Plot showing the norm of the function approximation error, $\|\varepsilon_i(\kappa_i)\|$, over time for all agents $i \in \mathcal{V}$. 104

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

DISTRIBUTED TARGET TRACKING IN NONLINEAR MULTI-AGENT SYSTEMS

By

Cristian F. Nino

December 2025

Chair: Warren E. Dixon
Major: Mechanical Engineering

This dissertation develops an assured control framework for multi-agent target tracking, addressing challenges of nonlinear dynamics, uncertainty, and constrained sensing. First, foundational methods are established, including an online ResNet-based adaptive controller for uncertain dynamics and an observer for target tracking using only relative position information. Second, a distributed, robust controller is developed for heterogeneous nonlinear systems. Using a Robust Integral of the Sign of the Error (RISE) methodology, a novel P-function construction enables a fully distributed, single-hop design that guarantees exponential tracking convergence. Third, an adaptive framework is designed for directed graphs with limited sensing. It uses an online deep residual network observer to learn unknown target dynamics, and a generalized trackability condition ensures stability using only partial, relative position measurements. Validated by constructive Lyapunov analysis, this integration of robust and adaptive strategies provides a rigorous foundation for assured performance in distributed, nonlinear multi-agent environments.

CHAPTER 1 INTRODUCTION

The study of multi-agent systems has expanded significantly, driven by applications from swarm robotics to communication networks. A primary challenge in modern systems is the integration of heterogeneous agents. These agents may operate in multiple domains, such as air, ground, and maritime environments. Furthermore, individual agents often possess distinct dynamical models, computational capabilities, and operational time-scales.

Achieving a collective, system-wide objective requires network control. This field governs the interactions between agents to ensure coherent behavior. Foundational problems in network control, such as consensus, synchronization, and formation control, establish the principles for coordinating group dynamics.

The multi-agent target-tracking problem serves as a comprehensive abstraction of these coordination challenges. In this problem, a network of agents must collectively monitor or pursue a target. This capability is essential for applications including environmental monitoring, surveillance, and search and rescue operations.

Developing effective solutions for target tracking presents significant technical difficulties. Many existing control methods rely on simplifying assumptions that are unrealistic in practical deployments. A common assumption is the availability of full state feedback, including position and velocity, for each agent. In practice, measurements such as velocity may be unavailable or excessively noisy. This disparity necessitates the use of relative sensing, where agents must operate using only local measurements from onboard sensors without access to a global coordinate frame.

Similarly, theoretical models for inter-agent communication often presume fixed, undirected, or strongly connected graphs. Real-world networks are characterized by directed, intermittent, or rapidly switching topologies. Agents must contend with asynchronous updates, communication delays, and packet dropouts caused by environmental factors or network faults.

The physical dynamics of the agents and the target are also frequently uncertain. Controllers must be robust to unmodeled dynamics and external disturbances. Agents must navigate complex

physical environments, requiring solutions that account for obstacle avoidance and spatial constraints beyond simple Euclidean settings. Some existing methods, particularly those based on potential fields, can suffer from local minima and coordination deadlocks.

The combination of these challenges highlights the need for new approaches. Developing controllers for such systems requires a focus on robustness to handle uncertainty and disturbances. It demands adaptivity to learn and compensate for unknown system parameters or environmental changes. The solutions must be distributed, relying only on local information and computation. Finally, these methods must provide assurance of performance, guaranteeing stability and task completion despite the complex, nonlinear dynamics and challenging operational conditions. This text establishes the foundation for developing such assured, robust, and adaptive strategies for collaborative autonomy.

1.1 Background

The canonical multi-agent control problem is that of consensus, a field whose fundamental goal is to design local control laws that enable a group of agents to agree upon a common value, such as position or velocity, using only information from their local neighbors (e.g., [65, 64, 63, 78, 45]). This foundational work established the core principles of distributed coordination and inspired extensive research into related problems (e.g., [80, 56, 81, 43, 5]).

Many of these coordination tasks are formulated within a leader-follower architecture, where one or more agents act as leaders, possessing information about the collective objective, while the follower agents must coordinate to track the leaders' states (e.g., [13, 20]). This paradigm extends naturally to the dynamic problem of multi-agent target tracking, where the target itself is treated as a dynamic, and often non-cooperative, leader. This problem is prominent due to its versatility in modeling real-world scenarios involving second-order systems like manipulator robots, aircraft, and spacecraft (e.g., [4]). The leader-follower paradigm also encompasses problems such as indirect regulation, where a small number of informed agents are used to influence the collective behavior of a much larger group (e.g., [46, 88, 60]). Certain adversarial formulations of these problems are analyzed using the tools of differential games (e.g., [89, 96, 28, 50]).

Much of the initial literature on these varied problems has relied on ideal conditions, such as linear agent dynamics and complete information, to establish feasibility and design control laws.

1.1.1 Challenges in Sensing and Communication

A fundamental distinction in multi-agent control is the architecture for information processing and decision-making. Centralized control relies on a single processing unit that gathers all information from every agent and computes a global solution. While potentially optimal, this approach suffers from a single point of failure and high communication overhead.

To overcome this, distributed control strategies were developed, where agents make decisions based on their own local information and information received from a limited set of neighbors over a communication network. This approach often assumes some knowledge of the network structure, such as the total number of agents, but does not require a central coordinator. A more restrictive case is decentralized control, where each agent operates in complete isolation, using only its own onboard sensor measurements without any inter-agent communication.

A significant body of research assumes the availability of full state information for each agent. The practical difficulty or impossibility of obtaining complete state measurements, such as velocity, has motivated the development of control strategies based on incomplete information (e.g. [1]). A common example is the reliance on relative velocity measurements, which may not be accessible or accurate in various robotics applications. This has motivated a desire to develop methods that rely solely on relative position information, allowing for on-board computation using sensors like cameras or ultrasonic sensors without a common reference frame. When global state information is impractical to obtain, relative sensing is a viable alternative, as many aerial, ground, and underwater platforms operate where GPS or external beacons are degraded or absent (e.g., [1, 2, 12, 4]).

Distributed observers are one such approach, where agents collaboratively estimate the state of the target or other agents using only local measurements and communication (e.g., [19]). Initial work addressed the multi-agent tracking problem where the tracked leader has unmeasurable velocity states, but assumed first-order follower dynamics [20]. While second-order dynamics

were later considered, relative velocity measurements were still assumed to be known. This limitation was addressed in [19, 24] by incorporating a linear second-order dynamic model for agents and constructing a distributed observer. Subsequent results further expanded upon these ideas to incorporate heterogeneous linear dynamics. However, these results do not account for nonlinearity in the dynamics. The work in [58] proposed a distributed observer for second-order agents that relies solely on relative positions, establishing exponential tracking under heterogeneous, nonlinear agent dynamics and indicating that velocity-free strategies are feasible in realistic sensing regimes.

Other works have focused on output feedback control, designing controllers that function with only partial state measurements (e.g., [32]). Research on systems with even more constrained sensing, such as formation tracking or localization problems relying on bearing-only or range-only measurements, has also become prominent, particularly for applications where global positioning is denied (e.g., [47, 85]).

The communication network topology also presents a major challenge. Early research often considered static, undirected communication graphs, where information flow is bidirectional and constant. Subsequent work addressed more realistic cases such as directed graphs, which better represent sensing and communication constraints, as well as time-varying or switched network topologies (e.g., [55]). Directed information flow is common in field robotics due to asymmetric links and limited broadcast ranges. For directed graphs that are only rooted, the graph Laplacian matrix is asymmetric, which complicates standard ensemble-level stability proofs that rely on symmetry. While consensus and formation control over directed graphs are well studied (e.g., [55, 54, 99, 101]), a further challenge arises when the target agent's states are only partially observable. This creates a need to understand how the asymmetric topology and limited information jointly affect the system's ability to track the target. There is limited work that integrates (i) partial target feedback, (ii) relative-only measurements, and (iii) online adaptive learning in a single, implementable framework for target tracking over directed graphs.

The assumption of continuous communication, which is energetically costly and can lead to network congestion, has motivated the study of intermittent communication (exchanging information at discrete instants) and the paired challenge of asynchronous communication (agents updating on individual clocks) (e.g., [79, 77, 61]). To manage network bandwidth even further, event-triggered and self-triggered control strategies have been developed to reduce the frequency of information exchange while preserving stability and performance, along with extensive analysis into other non-idealities such as communication delays (e.g., [66, 34, 8, 62, 100, 84]).

1.1.2 Control of Uncertain Nonlinear Systems

Many physical systems are governed by nonlinear dynamics, for which linear control methods are inadequate, and the added presence of unmodeled dynamics and external disturbances necessitates the use of robust and adaptive control methods.

Robust control techniques guarantee performance in the face of bounded uncertainties, ensuring that the tracking error remains within acceptable bounds despite disturbances. Among these techniques, the Robust Integral of the Sign of the Error (RISE) methodology has demonstrated significant potential by achieving exponential tracking error convergence with continuous control inputs despite time-varying disturbances [67, 72, 70]. The RISE approach incorporates integration of a signum term and a specific P-function construction in the Lyapunov analysis. Extending this P-function to multi-agent systems is challenging due to the ensemble-level analysis. The first application to multi-agent tracking in [33] required 2-hop communication. Other RISE-based works have addressed related but distinct problems, such as flocking [98] or stabilization [23]. This approach can be viewed as a form of implicit learning, as it provides instantaneous reactivity to attenuate disturbances but requires continuous feedback to function effectively.

In contrast, for systems with parametric uncertainty where the model is unknown, adaptive controllers serve as a form of explicit learning by estimating the unknown parameters online to achieve the tracking objective. Neural networks (NNs) are well established for approximating unstructured uncertainties in continuous functions over compact domains (e.g., [91, 22, 31, 35]).

The evolution of NN-based control has progressed from single-layer architectures with Lyapunov-based adaptation (e.g., [41, 42, 73]) to more complex deep neural network (DNN) implementations, motivated by numerous examples of improved function approximation efficiency (e.g., [40, 15, 82]). Early DNN approaches develop Lyapunov-based adaptive update laws for the output layer while the inner layers are updated either in an iterative offline manner as in [26] and [93], or using modular adaptive update laws [38]. Recent developments have established frameworks for real-time adaptation of all DNN layers (e.g., [68, 83]) for various DNN architectures, addressing issues in transient performance (e.g., [39]) and leveraging persistence of excitation (e.g., [69]).

Deep residual neural network (ResNet) architectures have emerged as particularly promising candidates for adaptive control. The ResNet architecture is popular because it addresses optimization challenges that arise with increasing network depth by introducing "skip connections" that create direct paths for information flow (e.g., [17, 97]). These connections help prevent the vanishing gradient problem by learning the "residual" difference, which simplifies optimization. Theoretical analyses have demonstrated their favorable optimization properties (e.g., smoother loss landscapes [44], absence of spurious local optima [16, 30], and stability of equilibria [57]) and universal approximation capabilities (e.g., [48, 94, 49]). Critical advancements include pre-activation shortcuts [18], which improve information flow, and conceptual similarities to DenseNets [25]. Recently, [71] introduced the first Lyapunov-based ResNet for online control, though it utilized the original architecture without pre-activation and required matching input-output dimensions. Adaptive NNs have also been explored for state estimation, such as in [59], where a Lyapunov-based DNN learns unknown target behavior online and a "trackability" condition was introduced for systems with partial information, though this work assumed an undirected communication graph.

This adaptive approach offers the potential for extrapolation and provides better behavior in situations where feedback may be temporarily lost or degraded, but is often ineffective at learning

explicit functions of time and requires a sufficient richness of data for online parameter convergence.

The combination of these two approaches is particularly effective (e.g., [86, 75, 11]). Robustness provides baseline stability against unstructured uncertainties, while adaptation learns structured, parametric uncertainties to improve performance. This synergy is critical in the multi-agent context, where uncertainties propagate throughout the network and agent interactions create complex, difficult-to-model emergent dynamics. The fusion of robust and adaptive techniques, validated through constructive Lyapunov stability analysis, provides a rigorous foundation for assured performance in these complex, distributed, and nonlinear environments.

1.2 Dissertation Scope

The preceding literature review highlights a significant gap. While foundational multi-agent problems are well understood under ideal conditions, practical multi-agent systems are governed by nonlinear dynamics and are subject to significant uncertainties and operational constraints. This dissertation develops a scalable and resilient distributed framework to address the core problem of target tracking for agents that are heterogeneous, uncertain, and possess limited sensing.

To build this framework, the dissertation is organized as follows.

Chapter 2 establishes the mathematical preliminaries used throughout the dissertation. This includes notation and the fundamentals of nonsmooth analysis and algebraic graph theory.

Chapter 3 introduces foundational work on online ResNet-based adaptive control and multi-agent target tracking of second-order systems. This work develops an online ResNet-based adaptive control method for the target tracking of a single agent with unknown dynamics. An observer is also developed to solve the second-order target tracking problem using only relative position information.

Chapter 4 tackles the challenge of robust performance against external disturbances and unmodeled dynamics. A distributed controller using a Robust Integral of the Sign of the Error (RISE) methodology is developed. This approach provides assured stability, guaranteeing

exponential convergence for heterogeneous systems. The design is notable for its novel Lyapunov-based P-function construction, which enables a fully distributed implementation requiring only local, single-hop communication.

Chapter 5 addresses the practical limitations of sensing and communication. An adaptive framework is developed for systems operating over directed communication graphs and using only partial, relative position feedback. This solution employs deep residual neural networks to learn unknown target dynamics online. A generalized trackability condition for directed networks is established, and the locally implementable adaptation law ensures convergence without requiring velocity measurements or symmetric communication.

Chapter 6 concludes the dissertation by discussing directions for future research. This future work focuses on generalizing the developed frameworks using sheaf-theoretic tools to solve more complex problems involving multi-target tracking, heterogeneous agent objectives, and spatiotemporal communication constraints.

By integrating these robust and adaptive techniques, this work provides a comprehensive, assured solution for distributed control in complex, uncertain, and nonlinear multi-agent environments.

CHAPTER 2 PRELIMINARIES

This chapter introduces the notation, mathematical foundations, and core technical concepts used throughout the dissertation.

2.1 Notation and Mathematical Conventions

This section establishes the notation and core mathematical frameworks used in this work.

2.1.1 General and Set Notation

Let $\mathbb{R}$ denote the real numbers, $\mathbb{Z}$ the integers, and $\mathbb{C}$ the complex numbers. Define $\mathbb{Z}_{>0} \triangleq \{1, 2, \ldots\}$ and $\mathbb{Z}_{\geq 0} \triangleq \{0, 1, 2, \ldots\}$. For $M \in \mathbb{Z}_{>0}$, set $[M] \triangleq \{1, 2, \ldots, M\}$. Let P be a set and $Q(x)$ be a predicate on P. The universal quantifier $(\forall x \in P)\, Q(x)$ states that $Q(x)$ holds for every $x \in P$. The existential quantifier $(\exists x \in P)\, Q(x)$ states that $Q(x)$ holds for at least one $x \in P$. The cardinality of a set S is $|S|$; if S is finite, $|S|$ is its number of elements. The empty set is $\varnothing$, and any set S, $\varnothing \subseteq S$ and $|\varnothing| = 0$.

For propositions A and B, $A \wedge B$ means both are true, $A \vee B$ means at least one is true, and $\neg A$ is "not A". Implication is written as $A \implies B$ and is logically equivalent to $\neg A \vee B$. The phrase "if and only if" (abbreviated iff) means both directions hold with $A \iff B$ meaning $A \implies B$, and $B \implies A$.

Set-builder notation $\{x \in P : Q(x)\}$ denotes the elements of P that satisfy $Q(x)$. For sets A, B, containment $A \subseteq B$ means $(\forall x \in A)\ x \in B$ and the proper containment $A \subset B$ means $A \subseteq B$ and $A \neq B$. The intersection and union are $A \cap B \triangleq \{x : x \in A \wedge x \in B\}$ and $A \cup B \triangleq \{x : x \in A \vee x \in B\}$, respectively.

Given some sets A and B, a set-valued map F from A to subsets of B is denoted by $F : A \rightrightarrows B$. The notation $\overline{\mathrm{co}}A$ denotes the closed convex hull of the set A. The closure of a set S, denoted by $\overline{S}$, is the union of S and its boundary points. A set is compact if it is closed and bounded, and a set is precompact if its closure is compact. Given any sets $A, B \subset \mathbb{R}$, the notation $A \leq B$ is used to state $a \leq b$ for all $a \in A$ and $b \in B$.

2.1.2 Vector and Matrix Notation

For $n \in \mathbb{Z}_{>0}$, let $\mathbf{1}_n \in \mathbb{R}^n$ be the column vector whose entries are all one, and let $\mathbf{0}_n \in \mathbb{R}^n$ be the column vector whose entries are all zero. For $m, n, p \in \mathbb{Z}_{>0}$, let $\mathbf{0}_{m\times n}$ be the $m \times n$ zero matrix, and let I_p denote the $p \times p$ identity matrix. For $v \in \mathbb{R}^m$, v_i denotes its i^{th} entry. For $M \in \mathbb{R}^{m\times n}$, M_{ij} denotes the j^{th} entry of the i^{th} column. For a set A and input x, the indicator function is $\mathbf{1}_A(x)$, with $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ otherwise.

For $r \in \mathbb{R}^n$, the Euclidean norm is $\|r\| \triangleq \sqrt{r^\top r}$. The open ball is $\mathbb{B}_\varepsilon(x) \triangleq \left\{y \in \mathbb{R}^d : \|y - x\| < \varepsilon\right\}$ for $x \in \mathbb{R}^d$ and $\varepsilon > 0$. For $D \subseteq \mathbb{R}^d$, the interior is $\text{int}(D) \triangleq \{x : \exists \varepsilon > 0 \text{ with } \mathbb{B}_\varepsilon(x) \subseteq D\}$, and the boundary is $\partial D \triangleq \left\{x : \forall \varepsilon > 0, \mathbb{B}_\varepsilon(x) \cap D \neq \varnothing \text{ and } \mathbb{B}_\varepsilon(x) \cap \left(\mathbb{R}^d \setminus D\right) \neq \varnothing\right\}$.

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be symmetric if $A = A^\top$. The minimum and maximum eigenvalues of a symmetric matrix A are denoted by $\lambda_{\min}(A) \in \mathbb{R}$ and $\lambda_{\max}(A) \in \mathbb{R}$, respectively. For $A \in \mathbb{R}^{m\times n}$, the spectral norm is $\|A\| \triangleq \sqrt{\lambda_{\max}(A^\top A)} = \sigma_{\max}\{A\}$. The spectrum $\Lambda(A)$ of a square matrix A is the set of its eigenvalues. For any eigenvalue $\lambda \in \Lambda(A) \subseteq \mathbb{C}$, $\text{Re}(\lambda)$ denotes its real part.

For a symmetric matrix $A \in \mathbb{R}^{n\times n}$, A is called positive definite (PD), positive semidefinite (PSD), negative definite (ND), or negative semidefinite (NSD) if $x^\top Ax > 0$, $x^\top Ax \geq 0$, $x^\top Ax < 0$, or $x^\top Ax \leq 0$ for all $x \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, respectively. Equivalently, write $A \succ 0$, $A \succeq 0$, $A \prec 0$, or $A \preceq 0$, respectively. For symmetric A, B of the same size, the Loewner partial order is defined by $A \succeq B \iff A - B \succeq 0$ and $A \succ B \iff A - B \succ 0$, with analogous meanings for $\preceq$ and $\prec$.

Given $N \in \mathbb{Z}_{>0}$ and vectors $\{x_1, \ldots, x_N\}$, define the stacked vector $[x_i]_{i\in[N]} \triangleq \begin{bmatrix} x_1^\top & \cdots & x_N^\top \end{bmatrix}^\top \in \mathbb{R}^{nN}$, where $x_i \in \mathbb{R}^n$ for all $i \in [N]$. Given maps $\{F_i\}_{i\in[N]}$ with $F_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$, define the stacked map $F \triangleq [F_i]_{i\in[N]} : \prod_{i\in[N]} \mathbb{R}^{n_i} \to \prod_{i\in[N]} \mathbb{R}^{m_i}$ by $F\left([x_i]_{i\in[N]}\right) \triangleq [F_i(x_i)]_{i\in[N]}$.

For $v \in \mathbb{R}^m$, let v_i denote the i^{th} entry of v for all $i \in [m]$. For $v \in \mathbb{R}^m$, $\text{diag}\{v\} \in \mathbb{R}^{m\times m}$ is the diagonal matrix with diagonal entries $v_1, \ldots, v_m$. For $H \in \mathbb{R}^{m\times n}$ with columns $\{h_i\}_{i\in[n]} \subset \mathbb{R}^m$, the vectorization operator is $\text{vec}(H) \triangleq \begin{bmatrix} h_1^\top & \cdots & h_n^\top \end{bmatrix}^\top \in \mathbb{R}^{mn}$. Given $A \in \mathbb{R}^{p\times q}$ and $B \in \mathbb{R}^{m\times n}$,

define the Kronecker product $A \otimes B \in \mathbb{R}^{pm \times qn}$ by $A \otimes B \triangleq \left[A_{ij}B\right]_{i=1,\ldots,p;j=1,\ldots,q}$. For $A \in \mathbb{R}^{p \times q}$, $B \in \mathbb{R}^{a \times r}$, and $C \in \mathbb{R}^{r \times s}$, the identity $\text{vec}\,(ABC) = (C^\top \otimes A)\,\text{vec}\,(B)$ holds, and hence $\frac{\partial}{\partial \text{vec}(B)}\text{vec}\,(ABC) = C^\top \otimes A$. For a matrix $A \in \mathbb{R}^{m \times n}$, the column space $\text{col}\,(A) \subseteq \mathbb{R}^m$ is the set of all linear combinations of its columns, and the row space $\text{row}\,(A) \subseteq \mathbb{R}^n$ is defined analogously. For any linear subspace S, $\dim\,(S)$ denotes the cardinality of any basis of S. The rank of A is $\text{rank}\,(A) = \dim\,(\text{col}\,(A)) = \dim\,(\text{row}\,(A))$. The kernel is $\ker\,(A) \triangleq \{x \in \mathbb{R}^n : Ax = \mathbf{0}_m\}$. Viewing $A \in \mathbb{R}^{m \times n}$ as a linear map $\mathbb{R}^n \to \mathbb{R}^m$, injectivity holds if and only if $\ker\,(A) = \{\mathbf{0}_n\}$, equivalently $\text{rank}\,(A) = n$ with $n \leq m$; surjectivity holds if and only if $\text{col}\,(A) = \mathbb{R}^m$, equivalently $\text{rank}\,(A) = m$ with $m \leq n$. A matrix A has full rank when $\text{rank}\,(A) = \min\{m, n\}$. If $m = n$, then A is invertible if and only if it has full rank.

The Moore-Penrose inverse (pseudoinverse) of $A \in \mathbb{R}^{m \times n}$ is denoted as $A^+ \in \mathbb{R}^{n \times m}$. In the full row rank (surjective) case, $A^+ = A^\top \left(AA^\top\right)^{-1}$ and $AA^+ = I_m$; in the full column rank (injective) case, $A^+ = \left(A^\top A\right)^{-1} A^\top$ and $A^+A = I_n$.

For $A \in \mathbb{R}^{m_A \times n_A}$ and $B \in \mathbb{R}^{m_B \times n_B}$, the block-diagonalization operator is $\text{blkdiag}\,\{A, B\} \triangleq \begin{bmatrix} A & \mathbf{0}_{m_A \times n_B} \\ \mathbf{0}_{m_B \times n_A} & B \end{bmatrix} \in \mathbb{R}^{(m_A+m_B)\times(n_A+n_B)}$. Given $A = \text{blkdiag}\,\{A_1, \ldots, A_N\}$ with $A_i \in \mathbb{R}^{m_i \times n_i}$ and $x = [x_i]_{i\in[N]}$ where $x_i \in \mathbb{R}^{n_i}$, $Ax = [A_i x_i]_{i\in[N]}$. For an indexed family $\left\{A_p\right\}_{p=1}^m$ of conformable matrices, the right-to-left product is $\overset{\curvearrowleft}{\prod_{p=a}^{m}} A_p \triangleq A_m A_{m-1} \ldots A_a$, for $1 \leq a \leq m$, and is defined to be I of appropriate size when $a > m$.

2.1.3 Function Spaces and Nonsmooth Analysis

Let $A \subseteq \mathbb{R}^d$ be equipped with Lebesgue measure μ, and let $B \cong \mathbb{R}^m$ be a finite-dimensional normed space with norm $\|\cdot\|$. Define $\|f\|_\infty \triangleq \inf\{M > 0 : \exists N \subseteq A \text{ with } \mu(N) = 0 \text{ and } \|f(x)\| \leq M \text{ for all } x \in A \backslash N\}$. Let $\mathcal{L}_\infty(A; B)$ be the collection of all measurable $f : A \to B$ for which $\|f\|_\infty < \infty$. For $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$, let $C(A, B)$ denote the set of continuous mappings $f : A \to B$. A function is of class C^k if it has k continuous derivatives.

The notation "a.e." (almost everywhere) means that a property P holds for all $x \in X \backslash N$, where $N \in \Sigma$ and $\mu(N) = 0$ in a measure space (X, Σ, μ). Consider a Lebesgue measurable and locally essentially bounded function $h : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$. The Filippov regularization of h is defined as

$$K[h](y,t) \triangleq \bigcap_{\delta>0} \bigcap_{\mu\mathbb{S}=0} \overline{\text{co}} h(\mathbb{B}_\delta(y) \backslash \mathbb{S}, t),$$

where $\bigcap_{\mu\mathbb{S}=0}$ denotes the intersection over all sets $\mathbb{S}$ of Lebesgue measure zero. A function $y : \mathcal{I}_y \to \mathbb{R}^n$ is called a Filippov solution of $\dot{y} = h(y,t)$ on the interval $\mathcal{I}_y \subseteq \mathbb{R}_{\geq 0}$, if y is absolutely continuous on $\mathcal{I}_y$, and is a solution to the differential inclusion $\dot{y} \overset{\text{a.e.}}{\in} K[h](y,t)$. Clarke's generalized gradient for a locally Lipschitz function $V : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is defined as

$$\partial V(x,t) \triangleq \overline{\text{co}} \{\lim \nabla V(x,t) : (x_i, t_i) \to (x,t), (x_i, t_i) \notin \Omega_V\},$$

where Ω_V denotes the set of measure zero wherever ∇V is not defined.

2.1.4 Algebraic Graph Theory

Let $N \in \mathbb{Z}_{>0}$ and $\mathcal{V} \triangleq [N] = \{1, \ldots, N\}$. Consider a static, weighted, directed graph (digraph) $G \triangleq (\mathcal{V}, E, w)$, where $E \subseteq \mathcal{V} \times \mathcal{V}$ contains only arcs between distinct vertices, and $w : E \to \mathbb{R}_{>0}$ assigns a positive weight $w(e)$ to every arc $e \in E$. The relation $i \to j$ means $(i,j) \in E$. The relation $i \leftrightarrow j$ means $(i,j) \in E$ and $(j,i) \in E$ with $i \neq j$.

For each $i \in \mathcal{V}$, the in-neighborhood is $\mathcal{N}_i \triangleq \{j \in \mathcal{V} : (j,i) \in E, j \neq i\}$. Define the adjacency matrix $A \in \mathbb{R}^{N \times N}$ by $A_{ij} \triangleq w((j,i)) \mathbf{1}_{\{(j,i) \in E\}}$ so that $A_{ij} > 0$ when $j \to i$. The in-degree and out-degree of i are $d_i^{\text{in}} \triangleq \sum_{j \in \mathcal{N}_i} A_{ij} = (A\mathbf{1}_N)_i$ and $d_i^{\text{out}} \triangleq \sum_{j:(i,j) \in E} w((i,j)) = \sum_{j=1}^N A_{ji}$. Let $D \triangleq \text{diag}\{A\mathbf{1}_N\}$ be the in-degree matrix and let the in-Laplacian be $L \triangleq D - A$.

A directed path $i \rightsquigarrow j$ is a finite sequence of distinct arcs $(\mathtt{v}_1, \mathtt{v}_2), \ldots, (\mathtt{v}_{m-1}, \mathtt{v}_m) \in E$ with $\mathtt{v}_1 = i$ and $\mathtt{v}_m = j$. The reachable set of j is $\mathcal{R}(j) \triangleq \{i \in \mathcal{V} : j \rightsquigarrow i\} \cup \{j\}$. A set $\mathcal{R} \subseteq \mathcal{V}$ is a reach if $\mathcal{R} = \mathcal{R}(i)$ for some i and there is no j such that $\mathcal{R}(i) \subset \mathcal{R}(j)$. The digraph G is rooted if there exists at least one $r \in \mathcal{V}$, termed a root, such that $r \rightsquigarrow j$ for every $j \in \mathcal{V}$.

A subset $\mathcal{W} \subseteq \mathcal{V}$ is strongly connected if $i \rightsquigarrow j$ and $j \rightsquigarrow i$ for all $i, j \in \mathcal{W}$. A strongly connected component (SCC) is a maximal strongly connected subset of $\mathcal{V}$.

The condensation $\mathfrak{C}(G)$ is the directed acyclic graph (DAG) whose vertices are the SCCs of G, with an arc $\mathcal{S}_1 \to \mathcal{S}_2$ when some $v \in \mathcal{S}_1$ and $w \in \mathcal{S}_2$ satisfy $(v, w) \in E$. An SCC $\mathcal{S}$ is a root SCC if its vertex in $\mathfrak{C}(G)$ has in-degree zero; equivalently, no arc enters $\mathcal{S}$ from $\mathcal{V} \backslash \mathcal{S}$. Since $\mathfrak{C}(G)$ is acyclic, at least one root SCC exists; if G is rooted, there is only one root SCC.

2.1.5 Deep Residual Neural Network Model

Consider a fully connected feedforward ResNet with $b \in \mathbb{Z}_{\geq 0}$ building blocks, input $x \in \mathbb{R}^{L_{\text{in}}}$, and output $y \in \mathbb{R}^{L_{\text{out}}}$, as shown in Fig. 2-1. For each block index $i \in \{0, \ldots, b\}$, let $k_i \in \mathbb{Z}_{>0}$ be the number of hidden layers in the i^{th} block. Let $\vartheta_i \in \mathbb{R}^{L_{i,0}}$ denote the block input (with $\vartheta_0 \triangleq x$ and $L_{0,0} \triangleq L_{\text{in}}$) and let $\theta_i \in \mathbb{R}^{\mathrm{p}_i}$ be the vector of parameters (weights and biases) associated with the i^{th} block.

For each block $i \in \{0, \ldots, b\}$, let $L_{i,j} \in \mathbb{Z}_{>0}$ denote the number of neurons in the j^{th} layer for $j \in \{0, \ldots, k_i + 1\}$. Furthermore, define the augmented dimension $L^a_{i,j} \triangleq L_{i,j} + 1$, for all $(i, j) \in \{0, \ldots, b\} \times \{0, \ldots, k_i\}$. Each block function $\Phi_i : \mathbb{R}^{L^a_{i,0}} \times \mathbb{R}^{\mathrm{p}_i} \to \mathbb{R}^{L_{i,k_i+1}}$ is a fully connected feedforward DNN, with $L_{i,k_i+1} \triangleq L_{\text{out}}$ for all $i \in \{0, \ldots, b\}$, where $\mathrm{p}_i \triangleq \sum_{j=0}^{k_i} L^a_{i,j} L_{i,j+1}$. For any input $v \in \mathbb{R}^{L^a_{i,j}}$, the DNN is defined recursively by

$$\varphi_{i,j}(v) \triangleq \begin{cases} V_{i,0}^\top v, & j = 0, \\ V_{i,j}^\top \phi_{i,j}\left(\varphi_{i,j-1}(v)\right), & j \in \{1, \ldots, k_i\}, \end{cases} \tag{2-1}$$

with $\Phi_i(v, \theta_i) = \varphi_{i,k_i}(v)$.

For each $j \in \{0, 1, \ldots, k_i\}$ the matrix $V_{i,j} \in \mathbb{R}^{L^a_{i,j} \times L_{i,j+1}}$ contains the weights and biases; in particular, if a layer has n (augmented) inputs and the subsequent layer has m nodes, then $V \in \mathbb{R}^{n \times m}$ is constructed so that its $(i, j)^{\text{th}}$ entry represents the weight from the i^{th} node of the input to the j^{th} node of the output, with the last row corresponding to the bias terms. For the DNN architecture described by (2-1), the vector of DNN weights of the i^{th} block is $\theta_i \triangleq \begin{bmatrix} \left(\operatorname{vec}\left(V_{i,0}\right)\right)^\top & \cdots & \left(\operatorname{vec}\left(V_{i,k_i}\right)\right)^\top \end{bmatrix}^\top \in \mathbb{R}^{\mathrm{p}_i}$. The activation function $\phi_{i,j} : \mathbb{R}^{L_{i,j}} \to \mathbb{R}^{L^a_{i,j}}$ is

given by $\phi_{i,j}\left(\varphi_{i,j-1}\right) = \left[\varsigma_{i,j}\left(\left(\varphi_{i,j-1}\right)_1\right) \quad \varsigma_{i,j}\left(\left(\varphi_{i,j-1}\right)_2\right) \quad \cdots \quad \varsigma_{i,j}\left(\left(\varphi_{i,j-1}\right)_{L_{i,j}}\right) \quad 1\right]^\top \in \mathbb{R}^{L^a_{i,j}}$ where $\left(\varphi_{i,j-1}\right)_\ell$ denotes the ℓ^{th} component of $\varphi_{i,j-1}$ for all $\ell \in \left[L^a_{i,j}\right]$. Here, $\varsigma_{i,j} : \mathbb{R} \to \mathbb{R}$ is the smooth activation function used by layer j of block i, and the appended 1 accounts for the bias term, for all $(i, j) \in \{0, \ldots, b\} \times \{0, \ldots, k_i\}$.

A pre-activation design is used so that before each block (except block 0), the output of the previous block is processed by an external activation function. Specifically, for each block $i \in \{1, \ldots, b\}$, define the pre-activation mapping $\psi_i : \mathbb{R}^{L_{i,k_i+1}} \to \mathbb{R}^{L^a_{i,k_i+1}}$ by $\psi_i\left(\vartheta_i\right) = \left[\varrho_i\left(\left(\vartheta_i\right)_1\right) \quad \varrho_i\left(\left(\vartheta_i\right)_2\right) \quad \cdots \quad \varrho_i\left(\left(\vartheta_i\right)_{L_{i,k_i+1}}\right) \quad 1\right]^\top \in \mathbb{R}^{L^a_{i,k_i+1}}$ where $\left(\vartheta_i\right)_\ell$ denotes the ℓ^{th} component of ϑ_i. Here, $\varrho_i : \mathbb{R} \to \mathbb{R}$ is the single element-wise smooth pre-activation function used by block i, and the appended 1 accounts for the bias term. The output of ψ_i serves as the input to block i and the residual connection is implemented by adding the current block output to the pre-activated output from the previous block. Hence, the ResNet recursion is defined by

$$\vartheta_{i+1} \triangleq \begin{cases} \Phi_0\left(\vartheta^a_0, \theta_0\right), & i = 0, \\ \vartheta_i + \Phi_i\left(\psi_i\left(\vartheta_i\right), \theta_i\right), & i \in \{1, \ldots, b\}, \end{cases} \tag{2-2}$$

with output $y \in \mathbb{R}^{L_{\text{out}}}$ and overall parameter vector $\Theta \triangleq \begin{bmatrix} \theta_0^\top & \cdots & \theta_b^\top \end{bmatrix}^\top \in \mathbb{R}^p$, where $p \triangleq \sum_{i=0}^b \mathrm{p}_i$ denotes the total number of ResNet parameters, and where $\vartheta^a_0 \triangleq \begin{bmatrix} \vartheta_0^\top & 1 \end{bmatrix}^\top \in \mathbb{R}^{L^a_{0,0}}$ denotes the augmented input to block 0. Therefore, the complete ResNet is represented as $\Psi : \mathbb{R}^{L_{\text{in}}} \times \mathbb{R}^p \to \mathbb{R}^{L_{\text{out}}}$ expressed as $\Psi\left(\vartheta, \Theta\right) = \vartheta_{b+1}$.

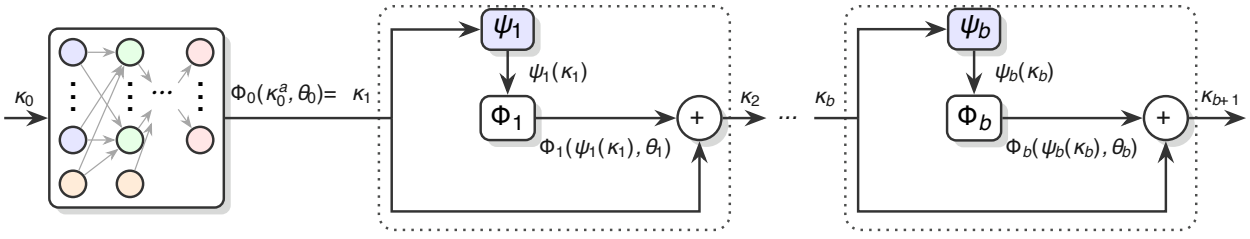

Figure 2-1. Deep Residual Neural Network Architecture.

The partial derivative of the ResNet with respect to its parameters is represented as $\frac{\partial}{\partial \Theta}\Psi\left(\vartheta,\Theta\right)=\left[\begin{array}{ccc}\frac{\partial}{\partial \theta_0}\Psi\left(\vartheta,\Theta\right) & \cdots & \frac{\partial}{\partial \theta_b}\Psi\left(\vartheta,\Theta\right)\end{array}\right]\in\mathbb{R}^{L_{\text{out}}\times p}$ and $\frac{\partial}{\partial \theta_i}\Psi\left(\vartheta,\Theta\right)=\left[\begin{array}{ccc}\frac{\partial}{\partial \text{vec}\left(V_{i,0}\right)}\Psi\left(\vartheta,\Theta\right) & \cdots & \frac{\partial}{\partial \text{vec}\left(V_{i,k_i}\right)}\Psi\left(\vartheta,\Theta\right)\end{array}\right]\in\mathbb{R}^{L_{\text{out}}\times \mathfrak{p}_i}$, where $\frac{\partial}{\partial \text{vec}\left(V_{i,j}\right)}\Psi\left(\vartheta,\Theta\right)\in\mathbb{R}^{L_{\text{out}}\times L_{i,j}^{a}L_{i,j+1}}$ for all $(i,j)\in\{0,\ldots,b\}\times\{0,\ldots,k_i\}$. Using (2-1), (2-2), and the property of the vectorization operator yields

$$\frac{\partial \Psi}{\partial \text{vec}\left(V_{i,j}\right)}=\left(\overset{\curvearrowleft}{\prod_{m=i+1}^{b}}\left(I_{L_{\text{out}}}+\left(\overset{\curvearrowleft}{\prod_{\ell=1}^{k_m}}V_{m,\ell}^{\top}\frac{\partial \phi_{m,\ell}}{\partial \varphi_{m,\ell-1}}\right)V_{m,0}^{\top}\frac{\partial \psi_m}{\partial \vartheta_m}\right)\right)\left(\overset{\curvearrowleft}{\prod_{\ell=j+1}^{k_i}}V_{i,\ell}^{\top}\frac{\partial \phi_{i,\ell}}{\partial \varphi_{i,\ell-1}}\right)\left(I_{L_{i,j+1}}\otimes \varkappa_{i,j}^{\top}\right),$$

where $\varkappa_{i,j}\triangleq\vartheta_0^a$ if $i=0$ and $j=0$, $\phi_{0,j}\left(\varphi_{0,j-1}\left(\vartheta_0^a\right)\right)$ if $i=0$ and $j>0$, $\psi_i\left(\vartheta_i\right)$ if $i>0$ and $j=0$, $\phi_{i,j}\left(\varphi_{i,j-1}\left(\psi_i\left(\vartheta_i\right)\right)\right)$ if $i>0$ and $j>0$.

For some $v\in\mathbb{R}^{L_{i,j}}$, the Jacobian $\frac{\partial \phi_{i,j}\left(\varphi_{i,j-1}(v)\right)}{\partial \varphi_{i,j-1}(v)}:\mathbb{R}^{L_{i,j}}\longrightarrow\mathbb{R}^{L_{i,j}^{a}\times L_{i,j}}$ of the activation function vector at the j^{th} layer is given by $\frac{\partial \phi_{i,j}\left(\varphi_{i,j-1}(v)\right)}{\partial \varphi_{i,j-1}(v)}=\left[\text{diag}\left\{\frac{\mathrm{d}\varsigma_{i,j}\left(\left(\varphi_{i,j-1}\right)_1\right)}{\mathrm{d}\left(\varphi_{i,j-1}\right)_1},\ldots,\frac{\mathrm{d}\varsigma_{i,j}\left(\left(\varphi_{i,j-1}\right)_{L_{i,j}}\right)}{\mathrm{d}\left(\varphi_{i,j-1}\right)_{L_{i,j}}}\right\}\mathbf{0}_{L_{i,j}}^{\top}\right]^{\top}$. Similarly, the Jacobian $\frac{\partial \psi_m(\vartheta_m)}{\partial \vartheta_m}:\mathbb{R}^{L_{\text{out}}}\longrightarrow\mathbb{R}^{L_{\text{out}}^{a}\times L_{\text{out}}}$ of the pre-activation function vector at block m is given by $\frac{\partial \psi_m(\vartheta_m)}{\partial \vartheta_m}=\left[\text{diag}\left\{\frac{\mathrm{d}\varrho_m\left((\vartheta_m)_1\right)}{\mathrm{d}(\vartheta_m)_1},\ldots,\frac{\mathrm{d}\varrho_m\left((\vartheta_m)_{L_{\text{out}}}\right)}{\mathrm{d}(\vartheta_m)_{L_{\text{out}}}}\right\}\mathbf{0}_{L_{\text{out}}}^{\top}\right]^{\top}$.

CHAPTER 3 FOUNDATIONAL METHODS

This chapter introduces the foundational methods that are extended upon in later chapters of the dissertation.

3.1 Online ResNet-Based Adaptive Control for Nonlinear Target Tracking

Neural networks are widely used to approximate unmodeled dynamics in control, with progress from single-layer adaptive schemes to deep models capable of real-time, all-layer adaptation. Residual networks (ResNets) are particularly suitable because skip connections mitigate vanishing gradients and simplify optimization, while theory indicates favorable loss landscapes and universal approximation. However, earlier Lyapunov-based ResNet controllers employed the original architecture without pre-activation shortcuts and enforced equal input-output dimensions, constraining learning efficiency and applicability. This work introduces a generalized ResNet for online nonlinear target tracking that places pre-activation shortcuts before nonlinearities to enhance information flow and adds a zeroth-layer block to accommodate dimension mismatches and strengthen feature propagation and reuse. A Lyapunov-based adaptation law updates all layers online and guarantees exponential convergence to a neighborhood of the target state under unknown dynamics and disturbances. Performance is validated through a comparative experiment, demonstrating improved tracking and broader applicability relative to prior ResNet-based controllers.

3.1.1 Problem Formulation

Consider the second-order nonlinear dynamical system described by the differential equation

$$\ddot{q} = f\left(q, \dot{q}\right) + g\left(q, \dot{q}, t\right) u\left(t\right) + \omega\left(t\right), \tag{3-1}$$

where $t_0 \in \mathbb{R}_{\geq 0}$ denotes the initial time, $t \geq t_0$ denotes the current time, $q : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the known generalized position, $\dot{q} : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the known generalized velocity, $\ddot{q} : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the unknown generalized acceleration, the unknown functions $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\omega : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ represent drift dynamics and exogenous disturbances,

respectively, $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0} \to \mathbb{R}^{n \times m}$ denotes a known control effectiveness matrix, $u : \mathbb{R}_{\geq t_0} \to \mathbb{R}^m$ denotes the control input, and $m \in \mathbb{Z}_{>0}$ denotes the number of control channels.

The reference trajectory is governed by the autonomous second-order system

$$\ddot{q}_d = f_d \left(q_d, \dot{q}_d \right), \tag{3-2}$$

where $q_d : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the known reference position, $\dot{q}_d : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the known reference velocity, $\ddot{q}_d : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the unknown reference acceleration, and $f_d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ represents the unknown reference dynamics.

Assumptions regarding the dynamical system outlined by (3-1) and (3-2) are provided as follows.

Assumption 3.1. *There exist known constants $\overline{q}_d, \overline{\dot{q}}_d \in \mathbb{R}_{>0}$ such that $\| q_d (t) \| \leq \overline{q}_d$ and $\| \dot{q}_d (t) \| \leq \overline{\dot{q}}_d$ for all $t \in \mathbb{R}_{\geq t_0}$.*

Assumption 3.2. *The functions f and f_d are of class C^1.*

Assumption 3.3. *The function g is of full row rank for all $(q, \dot{q}, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0}$ and is of class C^0. Furthermore, for each fixed $(q, \dot{q})$, the map $t \mapsto g (q, \dot{q}, t)$ is uniformly bounded in $t \in \mathbb{R}_{\geq t_0}$.*

Assumption 3.4. *The function ω is of class C^0 and there exists a known constant $\overline{\omega} \in \mathbb{R}_{\geq 0}$ such that $\| \omega (t) \| \leq \overline{\omega}$ for all $t \in \mathbb{R}_{\geq t_0}$.*

By [74], Assumption 3.3 ensures the the pseudoinverse $g^+ : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0} \to \mathbb{R}^{m \times n}$ acts as a right inverse for all $t \in \mathbb{R}_{\geq t_0}$. In addition, the following condition is imposed.

Assumption 3.5. *The function g^+ is of class C^0 and for each fixed $(q, \dot{q})$, the map $t \mapsto g^+ (q, \dot{q}, t)$ is uniformly bounded in $t \in \mathbb{R}_{\geq t_0}$.*

The control objective is to design a ResNet-based adaptive controller such that the state trajectory q is exponentially regulated to a neighborhood of the reference trajectory q_d, despite the

presence of unknown dynamics and bounded disturbances. To facilitate the control objective, define the tracking error $e \in \mathbb{R}^n$ as

$$e \triangleq q_d - q. \tag{3-3}$$

3.1.1.1 Control Synthesis

To facilitate the control design, the auxiliary tracking error function $r \in \mathbb{R}^n$ is defined as

$$r \triangleq \dot{e} + k_1 e, \tag{3-4}$$

where $k_1 \in \mathbb{R}_{>0}$ is a constant control gain. Differentiating (3-4) with respect to time and substituting (3-1)-(3-4) yields

$$\begin{aligned}
\dot{r} &= \ddot{e} + k_1 \dot{e} \\
&= \ddot{q}_d - \ddot{q} + k_1 \left(r - k_1 e \right) \\
&= f_d \left(q_d, \dot{q}_d \right) - \left(f \left(q, \dot{q} \right) + g \left(q, \dot{q}, t \right) u + \omega \left(t \right) \right) + k_1 \left(r - k_1 e \right) \\
&= f_d \left(q_d, \dot{q}_d \right) - f \left(q, \dot{q} \right) - g \left(q, \dot{q}, t \right) u - \omega \left(t \right) + k_1 \left(r - k_1 e \right) \\
&= h \left(q, \dot{q}, q_d, \dot{q}_d \right) - g \left(q, \dot{q}, t \right) u - \omega \left(t \right) + k_1 \left(r - k_1 e \right)
\end{aligned} \tag{3-5}$$

where $h : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is defined as $h \left(q, \dot{q}, q_d, \dot{q}_d \right) \triangleq f_d \left(q_d, \dot{q}_d \right) - f \left(q, \dot{q} \right)$.

3.1.1.2 Residual Neural Network Function Approximation

The ResNet architecture, characterized by skip connections and hierarchical feature extraction, models incremental changes rather than complete transformations of the underlying nonlinear mapping. This architecture learns the differences (or "residuals") between input and desired output at each layer, thereby enabling effective function approximation for complex nonlinear systems without requiring explicit governing equations.

To approximate the unknown dynamics given by $h \left(q, \dot{q}, q_d, \dot{q}_d \right)$ in (3-5), define the ResNet input vector $\kappa \in \mathbb{R}^{4n}$ as $\kappa \triangleq \begin{bmatrix} q^\top & \dot{q}^\top & q_d^\top & \dot{q}_d^\top \end{bmatrix}^\top \in \Omega$, where $\Omega \subset \mathbb{R}^{4n}$ is a compact set over which the universal approximation property holds. The ResNet-based approximation of $h \left(\kappa \right)$ is

given by $\Psi\left(\kappa, \widehat{\Theta}\right)$, where $\Psi : \mathbb{R}^{4n} \times \mathbb{R}^p \to \mathbb{R}^n$ denotes the ResNet architecture mapping and $\widehat{\Theta} \in \mathbb{R}^p$ denotes the adaptive parameter estimate. By Assumption 3.2, h is continuous on Ω.

Fix a user-prescribed accuracy $\overline{\varepsilon} > 0$. By the universal approximation property on the compact domain Ω (e.g., [94, Corollary 5.2]), there exists a corresponding ResNet architecture (hence, $p \in \mathbb{Z}_{>0}$) and a compact parameter search space $\mathfrak{V} \in \mathbb{R}^p$ with C^∞ boundary and $\mathbf{0}_p \in \text{int}\,(\mathfrak{V})$ such that $\inf_{\vartheta \in \mathfrak{V}} \sup_{\kappa \in \Omega} \|h\,(\kappa) - \Psi\,(\kappa, \vartheta)\| \leq \overline{\varepsilon}$.

For subsequent analysis, introduce $\Theta^* \in \arg\min_{\vartheta \in \mathfrak{V}} \sup_{\kappa \in \Omega} \|h\,(\kappa) - \Psi\,(\kappa, \vartheta)\|$ under the fixed $\Psi\,(\cdot, \cdot)$. By joint continuity of $(\kappa, \vartheta) \mapsto \|h\,(\kappa) - \Psi\,(\kappa, \vartheta)\|$ on $\Omega \times \mathfrak{V}$ and compactness of Ω, the map $\vartheta \mapsto \sup_{\kappa \in \Omega} \|h\,(\kappa) - \Psi\,(\kappa, \vartheta)\|$ is continuous. Since $\mathfrak{V}$ is compact, a minimizer exists by Weierstrass's Theorem (e.g., [36, Theorem 10.55]). Fix such a Θ^*. With this choice, $\sup_{\kappa \in \Omega} \|h\,(\kappa) - \Psi\,(\kappa, \Theta^*)\| \leq \overline{\varepsilon}$. Using this fixed parameter, write

$$h\,(\kappa) = \Psi\,(\kappa, \Theta^*) + \varepsilon\,(\kappa)\,, \tag{3-6}$$

where $\varepsilon\,(\kappa) \triangleq h\,(\kappa) - \Psi\,(\kappa, \Theta^*)$ satisfies the uniform bound $\sup_{\kappa \in \Omega} \|\varepsilon\,(\kappa)\| \leq \overline{\varepsilon}$. To see this, observe that since h and $\kappa \mapsto \Psi\,(\kappa, \theta^*)$ are continuous on Ω, ε is continuous on Ω. Because the Euclidean norm is continuous, $\kappa \mapsto \|\varepsilon\,(\kappa)\|$ is continuous on Ω. By Weierstrass's Theorem, $\varepsilon\,(\cdot)$ attains a finite maximum on Ω (e.g., [36, Theorem 10.55]). Define $\overline{\varepsilon}_{\text{act}} \triangleq \max_{\xi \in \Omega} \|\varepsilon\,(\xi)\|$ which satisfies $\overline{\varepsilon}_{\text{act}} \leq \overline{\varepsilon}$ by construction.

The same compactness-and-continuity argument does not rely on the specific ResNet notation and extends verbatim to any continuous target and any jointly continuous architecture with a compact parameter set containing the origin in its interior.

Remark 3.1. Let $\Omega \subset \mathbb{R}^m$ be nonempty and compact. Let $f : \Omega \to \mathbb{R}^n$ be continuous and let $\Phi : \Omega \times \mathbb{R}^p \to \mathbb{R}^n$ be jointly continuous. Fix $\varepsilon > 0$. Define

$$\mathcal{F}_\varepsilon \triangleq \left\{(\Phi, \mathfrak{V}) : \inf_{\theta \in \mathfrak{V}} \sup_{x \in \Omega} \|f\,(x) - \Phi\,(x, \theta)\| \leq \varepsilon\right\},$$

where $\mho \subset \mathbb{R}^p$ is compact, has C^∞ boundary, and satisfies $\mathbf{0}_p \in \operatorname{int}(\mho)$. The set $\mathcal{F}_\varepsilon$ is nonempty. Indeed, by universal approximation on Ω, there exists an architecture Φ and a parameter θ_ε whose uniform error does not exceed ε. Choosing any $R > \|\theta_\varepsilon\|$ and setting $\mho = \{\theta : \|\theta\| \leq R\}$ produces a pair $(\Phi, \mho)$ in $\mathcal{F}_\varepsilon$.

Fix any $(\Phi, \mho) \in \mathcal{F}_\varepsilon$ and consider the optimization problem

$$\min_{\theta \in \mho} \sup_{x \in \Omega} \|f(x) - \Phi(x, \theta)\|.$$

Define the corresponding minimizer set

$$\Theta \triangleq \arg\min_{\theta \in \mho} \sup_{x \in \Omega} \|f(x) - \Phi(x, \theta)\|.$$

Joint continuity of Φ and compactness of $\mho$ imply that Θ is nonempty (e.g., [36, Theorem 10.55]). Select any $\theta^* \in \Theta$ and define $\Phi^*(x) \triangleq \Phi(x, \theta^*)$. Then

$$\sup_{x \in \Omega} \|f(x) - \Phi^*(x)\| \leq \varepsilon.$$

Thus, after fixing ε, one may first choose a feasible architecture–parameter-set pair $(\Phi, \mho)$ and then choose a particular parameter $\theta^* \in \mho$ that achieves error at most ε on Ω. This is the same construction used in Section 3.1.1.2, where f corresponds to the unknown dynamics, Φ corresponds to the ResNet mapping $\Psi(\cdot, \cdot)$, the set $\mho$ is the compact parameter search space containing the origin in its interior, and θ^* is the selected parameter that realizes the approximation error bound.

Remark 3.2. The discussion above hides an implicit feasibility requirement linking the target accuracy ε, the allowed parameter radius, and the chosen architecture Φ. To make this explicit, fix an architecture Φ and a radius $R > 0$, and define

$$\mho(R) \triangleq \{\theta : \|\theta\| \leq R\}, \quad E_\Phi(R) \triangleq \inf_{\|\theta\| \leq R} \sup_{x \in \Omega} \|f(x) - \Phi(x, \theta)\|.$$

The quantity $E_\Phi(R)$ is the best uniform error that Φ can achieve on Ω using parameters whose norm does not exceed R. The map $R \mapsto E_\Phi(R)$ is nonincreasing. The pair $(\Phi, \mho(R))$ belongs to $\mathcal{F}_\varepsilon$ if and only if $E_\Phi(R) \le \varepsilon$. In particular, if the desired accuracy ε is chosen too small relative to the allowed radius R (or relative to the expressive power of Φ), then no parameter θ with $\|\theta\| \le R$ will satisfy the uniform error requirement. This does not contradict the construction above; it only means that such a pair $(\Phi, \mho(R))$ cannot be selected at the stage where $(\Phi, \mho) \in \mathcal{F}_\varepsilon$ is chosen.

A simple example illustrates this point. Take $f(x) = x$ on $\Omega = [-1, 1]$ and $\Phi(x, \omega) = \omega x$ with $|\omega| \le R$. Then

$$\sup_{|x|\le 1} |f(x) - \Phi(x, \omega)| = \sup_{|x|\le 1} |(1 - \omega) x| = |1 - \omega| .$$

Minimizing over $|\omega| \le R$ gives $\omega^* = \mathrm{proj}_{[-R,R]}(1)$, where $\mathrm{proj}_{[a,b]}(y) \triangleq \min\{\max\{y, a\}, b\}$ denotes projection onto $[a, b]$. Hence,

$$E_\Phi(R) = \inf_{|\omega|\le R} |1 - \omega| = \max\{0, 1 - R\} .$$

If $R < 1$ and $\varepsilon < 1 - R$, then $E_\Phi(R) > \varepsilon$. In that case, the ball $\mho(R) = \{\omega : |\omega| \le R\}$ is too small to reach accuracy ε, so $(\Phi, \mho(R)) \notin \mathcal{F}_\varepsilon$. Increasing the allowable radius R, or enlarging the architecture class so that Φ is more expressive, restores feasibility.

In practice, the supremum over Ω is not evaluated exactly but is approximated by the maximum error over a finite set of sampled states in Ω. The informal design loop is therefore: first, select a desired accuracy ε; next, increase R (and, if necessary, the size of the architecture Φ) until training with the norm constraint $\|\theta\| \le R$ achieves a maximum observed error that is at most ε; finally, define $\mho(R) \triangleq \{\theta : \|\theta\| \le R\}$. At that point, $\mho$ contains the corresponding minimizing parameter, $\mathbf{0}_p$ lies in its interior, and $(\Phi, \mho) \in \mathcal{F}_\varepsilon$ in the sense described above.

3.1.1.3 Control Design

Based on (3-5) and the subsequent stability analysis, the control input is designed as

$$u = g^{+}\left(q, \dot{q}, t\right)\left(\left(1 - k_1^2\right)e + \left(k_1 + k_2\right)r + \Psi\left(\kappa, \widehat{\Theta}\right)\right) \quad , \tag{3-7}$$

where $k_2 \in \mathbb{R}_{>0}$ is a constant control gain. Substituting (3-6) and (3-7) into (3-5) yields

$$\dot{r} = -k_2 r - e + \varepsilon\left(\kappa\right) - \omega\left(t\right) + \Psi\left(\kappa, \Theta^{*}\right) - \Psi\left(\kappa, \widehat{\Theta}\right). \tag{3-8}$$

Based on the subsequent stability analysis, the adaptive update law for $\widehat{\Theta}$ is designed as

$$\dot{\widehat{\Theta}} = \text{proj}\left(\Gamma\left(\frac{\partial \Psi\left(\kappa, \widehat{\Theta}\right)}{\partial \widehat{\Theta}}^{\top} r - k_3 \widehat{\Theta}\right)\right), \tag{3-9}$$

where $k_3 \in \mathbb{R}_{>0}$ is a user-defined forgetting rate, $\Gamma \in \mathbb{R}^{p \times p}$ is a user-defined symmetric positive-definite learning rate matrix, and the projection operator is defined as in [52, Appendix E, Equation E.4] and ensures the parameter estimates satisfy $\left\|\widehat{\Theta}\left(t\right)\right\| \leq \overline{\Theta}$ for all $t \in \mathbb{R}_{\geq t_0}$, where $\overline{\Theta} \triangleq \sqrt{\Theta_{\max}^2 + \overline{\epsilon}}$, $\Theta_{\max} \triangleq \sup_{\vartheta \in \mathcal{U}} \|\vartheta\|$, and $\overline{\epsilon} > 0$ is a margin introduced to ensure Lipschitz continuity of the projection mapping.

3.1.2 Stability Analysis

The ResNet mapping $\Psi\left(\kappa, \Theta\right)$ described in (3-6) is inherently nonlinear with respect to its weights. Designing adaptive controllers and performing stability analyses for systems that are nonlinearly parameterizable presents significant theoretical challenges. To address this nonlinearity, a first-order application of Taylor's theorem about the current parameter estimate is used. To quantify the approximation, the parameter estimation error $\widetilde{\Theta} \in \mathbb{R}^p$ is defined as

$$\widetilde{\Theta} \triangleq \Theta^{*} - \widehat{\Theta}. \tag{3-10}$$

Applying first-order Taylor's theorem to $\Psi(\kappa, \Theta^*)$ about the point $\left(\kappa, \widehat{\Theta}\right)$, and using (3-10) yields

$$\Psi(\kappa, \Theta^*) = \Psi\left(\kappa, \widehat{\Theta}\right) + \frac{\partial}{\partial \widehat{\Theta}} \Psi\left(\kappa, \widehat{\Theta}\right) \widetilde{\Theta} + R\left(\kappa, \widetilde{\Theta}\right), \tag{3-11}$$

where $\frac{\partial}{\partial \widehat{\Theta}} \Psi\left(\kappa, \widehat{\Theta}\right) \triangleq \left.\frac{\partial}{\partial \vartheta} \Psi(\kappa, \vartheta)\right|_{\widehat{\Theta}} \in \mathbb{R}^{n \times p}$ and $R : \mathbb{R}^{4n} \times \mathbb{R}^p \to \mathbb{R}^n$ denotes the first Lagrange remainder. Substituting (3-11) into (3-8) yields

$$\dot{r} = \frac{\partial}{\partial \widehat{\Theta}} \Psi\left(\kappa, \widehat{\Theta}\right) \widetilde{\Theta} + R\left(\kappa, \widetilde{\Theta}\right) + \varepsilon(\kappa) - k_2 r - e - \omega(t). \tag{3-12}$$

Let $z \in \mathbb{R}^\varphi$ denote the concatenated state vector $z \triangleq \begin{bmatrix} e^\top & r^\top & \widetilde{\Theta}^\top \end{bmatrix}^\top$, where $\varphi \triangleq 2n + p$. The evolution of z is governed by the initial value problem

$$\dot{z} = \mathtt{f}(z, t), \ z(t_0) = z_0, \tag{3-13}$$

$z_0 \in \mathbb{R}^\varphi$ is the initial state and, using (3-3), (3-9), (3-10), and (3-12), the vector field $\mathtt{f} : \mathbb{R}^\varphi \times \mathbb{R}_{\geq 0} \to \mathbb{R}^\varphi$ is defined as

$$\mathtt{f}(z, t) = \begin{bmatrix} r - k_1 e \\ \frac{\partial}{\partial \widehat{\Theta}} \Psi\left(\kappa, \widehat{\Theta}\right) \widetilde{\Theta} + R\left(\kappa, \widetilde{\Theta}\right) + \varepsilon(\kappa) - k_2 r - e - \omega(t). \\ -\mathrm{proj}\left(\Gamma\left(\left(\frac{\partial}{\partial \widehat{\Theta}} \Psi\left(\kappa, \widehat{\Theta}\right)\right)^\top r - k_3 \widehat{\Theta}\right)\right) \end{bmatrix}. \tag{3-14}$$

Following Assumptions 3.2 and 3.4, and invoking [52, Lemma E.1], $\mathtt{f}$ is locally Lipschitz in z and continuous in t. Therefore, by the Picard-Lindelöf theorem (e.g., [7, Chapter 1, Theorem 3.1]), for every $z_0 \in \mathbb{R}^\varphi$ there exists a unique maximal solution.

Since the universal approximation property of the ResNet holds only on the compact domain Ω, the subsequent stability analysis requires ensuring $\kappa \in \Omega$ for all $t \geq t_0$. This outcome is achieved by yielding a stability result which constrains z to a compact domain. Consider the

Lyapunov function candidate $V : \mathbb{R}^{\varphi} \to \mathbb{R}_{\geq 0}$ defined as

$$V(z) \triangleq \frac{1}{2} z^{\top} P z, \tag{3-15}$$

where $P \triangleq \text{blkdiag}\left\{I_{2n}, \Gamma^{-1}\right\} \in \mathbb{R}^{\varphi \times \varphi}$. By the Rayleigh quotient, (3-15) satisfies

$$\frac{1}{2}\lambda_1 \|z\|^2 \leq V(z) \leq \frac{1}{2}\lambda_{\varphi} \|z\|^2, \tag{3-16}$$

where $\lambda_1 \triangleq \lambda_{\min}\{P\} = \min\left\{1, \lambda_{\min}\left\{\Gamma^{-1}\right\}\right\}$ and $\lambda_{\varphi} \triangleq \lambda_{\max}\{P\} = \max\left\{1, \lambda_{\max}\left\{\Gamma^{-1}\right\}\right\}$.

By the mean value theorem (e.g., [37, Theorem 3.9]) and Taylor's theorem (e.g., [37, Theorem 4.7]), there exists a polynomial $\rho_0(\|\kappa\|) = a_2\|\kappa\|^2 + a_1\|\kappa\| + a_0$ with $a_2, a_1, a_0 \in \mathbb{R}_{>0}$ such that $\left\|R\left(\kappa, \widetilde{\Theta}\right)\right\| \leq \rho_0(\|\kappa\|)\left\|\widetilde{\Theta}\right\|^2$. Using the fundamental theorem of calculus and Assumption 3.2, there exists a function $\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined as $\rho(\|z\|) \triangleq 3\overline{\Theta}^2 \rho_0^2\left((k_1+2)\|z\| + 2\left(\overline{q}_d + \overline{\dot{q}}_d\right)\right)$ which satisfies $\overline{\rho}(\cdot) \triangleq \rho(\cdot) - \rho(0) \in \mathcal{K}_{\infty}$. Based on the subsequent analysis, define $\delta \triangleq \frac{3(\overline{\omega}+\overline{\varepsilon})^2}{4k_2} + \frac{k_3\overline{\Theta}^2}{2}$ and $k_{\min} \triangleq \min\left\{k_1, \frac{1}{3}k_2, \frac{1}{2}k_3\right\}$. The region to which the state trajectory is constrained is defined as

$$\mathcal{D} \triangleq \left\{\iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq \overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho(0)\right)\right\}, \tag{3-17}$$

where $\lambda_V \in \mathbb{R}_{>0}$ is a user-defined parameter which controls the desired rate of convergence. The compact domain over which the universal approximation property must hold is selected as

$$\Omega \triangleq \left\{\iota \in \mathbb{R}^{4n} : \|\iota\| \leq 2\left(\overline{q}_d + \overline{\dot{q}}_d\right) + (k_1+2)\,\overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho(0)\right)\right\}. \tag{3-18}$$

For the dynamical system described by (3-13), the set of initial conditions is defined as

$$\mathcal{S} \triangleq \left\{\iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq \sqrt{\frac{\lambda_1}{\lambda_{\varphi}}}\overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho(0)\right) - \sqrt{\frac{\delta}{\lambda_V}}\right\}, \tag{3-19}$$

and the uniform ultimate bound is defined as

$$\mathcal{U} \triangleq \left\{ \iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq \sqrt{\frac{\lambda_{\varphi}\delta}{\lambda_1 \lambda_V}} \right\}. \tag{3-20}$$

Theorem 3.1. *Consider the dynamical system described by (3-1) and (3-2). For any $z_0 \in \mathcal{S}$, the controller given by (3-7) and the adaptation law given by (3-9) ensure that z uniformly exponentially converges to $\mathcal{U}$ with estimate*

$$\|z(t)\| \leq \sqrt{\frac{\lambda_{\varphi}}{\lambda_1}\|z_0\|^2 \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}(t-t_0)} + \frac{\lambda_{\varphi}\delta}{\lambda_1\lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}(t-t_0)}\right)},$$

for all $t \in \mathbb{R}_{\geq t_0}$, provided that the sufficient gain condition $k_{\min} > \lambda_V + \frac{1}{k_2}\rho\left(\sqrt{\frac{\lambda_{\varphi}\delta}{\lambda_1\lambda_V}}\right)$ is satisfied and Assumptions 3.1-3.4 hold.

Proof. Taking the total derivative of (3-15) along the trajectories of (3-13) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}V(z(t)) = e^{\top}(t)\,\dot{e}(t) + r^{\top}(t)\,\dot{r}(t) + \widetilde{\Theta}^{\top}(t)\,\Gamma^{-1}\dot{\widetilde{\Theta}}(t) \quad . \tag{3-21}$$

Substituting (3-14), invoking [52, Appendix E.4], and using (3-10) yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V(z(t)) \leq{}& -k_1\|e(t)\|^2 - k_2\|r(t)\|^2 - k_3\left\|\widetilde{\Theta}(t)\right\|^2 \\ &+ r^{\top}(t)\left(R\left(\kappa(t), \widetilde{\Theta}(t)\right) + \varepsilon(\kappa(t))\right) - r^{\top}(t)\,\omega(t) + k_3\widetilde{\Theta}^{\top}(t)\,\Theta^*.\end{aligned} \tag{3-22}$$

Using (3-3), Assumption 3.1, and the triangle inequality yields $\|q(t)\| \leq \|e(t)\| + \overline{q}_d$. Similarly, using (3-4), Assumption 3.1, and the triangle inequality yields $\|\dot{q}(t)\| \leq k_1\|e(t)\| + \|r(t)\| + \overline{\dot{q}}_d$. Therefore, using the definition of κ yields

$$\|\kappa\| \leq (k_1 + 2)\|z\| + 2\left(\overline{q}_d + \overline{\dot{q}}_d\right). \tag{3-23}$$

Using (3-10) and (3-23), and the definition of $\overline{\Theta}$ yields

$$\left\| R\left(\kappa, \widetilde{\Theta}\right) + \varepsilon\left(\kappa\right) \right\| \leq 2\overline{\Theta}\rho_0 \left(\left(k_1 + 2\right) \|z\| + 2\left(\overline{q}_d + \overline{\dot{q}}_d\right) \right) \left\| \widetilde{\Theta} \right\| + \overline{\varepsilon}.$$

Thus, using Young's inequality and the definitions of δ and $k_{\min}$ yields that (3-22) is bounded above as

$$\frac{\mathrm{d}}{\mathrm{d}t} V\left(z\left(t\right)\right) \leq -\left(k_{\min} - \frac{\rho\left(\|z\left(t\right)\|\right)}{k_2} \right) \|z\left(t\right)\|^2 + \delta. \tag{3-24}$$

It is now shown, by contradiction, that the maximal solution to (3-13) exists for all time and remains in the interior of $\mathcal{D}$. Fix any $z_0 \in \mathcal{S}$ and let $z : [t_0, T_{\max}) \to \mathbb{R}^{\varphi}$ be the corresponding unique maximal solution of (3-13). Let $\mathcal{I} \triangleq \{t \in [t_0, T_{\max}) : z\left(\tau\right) \in \mathcal{D} \text{ for all } \tau \in [t_0, t]\}$, where $T_{\max} \in (t_0, \infty]$ denotes the supremum of all times $T > t_0$ for which a solution exists on $[t_0, T)$. Since $z_0 \in \mathcal{S} \subset \mathcal{D}$ and $t \mapsto z\left(t\right)$ is continuous, the set $\mathcal{I}$ is non-empty and contains an interval $[t_0, t_0 + \epsilon)$ for some $\epsilon > 0$.

For any $t \in \mathcal{I}$, the construction of $\mathcal{D}$ gives $\overline{\rho}\left(\|z\left(t\right)\|\right) \leq k_2\left(k_{\min} - \lambda_V\right) - \rho\left(0\right)$. Hence, using $\overline{\rho}\left(\cdot\right) = \varrho\left(\cdot\right) - \varrho\left(0\right)$ implies $k_{\min} - \frac{\rho(\|z(t)\|)}{k_2} \geq \lambda_V$. Therefore, combining (3-16) with (3-24) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} V\left(z\left(t\right)\right) \leq -\frac{2\lambda_V}{\lambda_{\varphi}} V\left(z\left(t\right)\right) + \delta, \tag{3-25}$$

for all $t \in \mathcal{I}$. Applying the Grönwall-Bellman inequality to (3-25) yields

$$V\left(z\left(t\right)\right) \leq V\left(z\left(t_0\right)\right) \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}\left(t - t_0\right)} + \frac{\lambda_{\varphi}\delta}{2\lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}\left(t - t_0\right)}\right), \tag{3-26}$$

for all $t \in \mathcal{I}$. Applying (3-16) to (3-26) yields

$$\|z\left(t\right)\| \leq \sqrt{\frac{\lambda_{\varphi}}{\lambda_1} \|z\left(t_0\right)\|^2 \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}\left(t - t_0\right)} + \frac{\lambda_{\varphi}\delta}{\lambda_1\lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_{\varphi}}\left(t - t_0\right)}\right)}, \tag{3-27}$$

for all $t \in \mathcal{I}$. Using (3-19), since $z_0 \in \mathcal{S}$, $\|z_0\| \leq \sqrt{\frac{\lambda_1}{\lambda_\varphi}}\overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho\left(0\right)\right) - \sqrt{\frac{\delta}{\lambda_V}}$. Substituting this expression into (3-27) and using the facts that $\mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} \leq 1$ and $1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} < 1$ for all $t \geq t_0$ yield the strict bound $\|z\left(t\right)\| < \overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho\left(0\right)\right)$ for all $t \in \mathcal{I}$. Thus, by (3-17), $z\left(t\right) \in \operatorname{int}\left(\mathcal{D}\right)$ for all $t \in \mathcal{I}$. Now, assume for contradiction that the maximal time of existence is finite, i.e., $\sup\left(\mathcal{I}\right) < T_{\max}$. If $\sup\left(\mathcal{I}\right) < T_{\max}$, continuity of $z\left(\cdot\right)$ would imply $z\left(\sup\left(\mathcal{I}\right)\right) \in \partial\mathcal{D}$, contradicting $z\left(t\right) \in \operatorname{int}\left(\mathcal{D}\right)$ for all $t \in \mathcal{I}$. Hence, $\sup\left(\mathcal{I}\right) = T_{\max}$ and $\mathcal{I} = [t_0, T_{\max})$, i.e., $z\left(t\right) \in \mathcal{D}$ for all $t \in [t_0, T_{\max})$.

Since $z\left(t\right) \in \mathcal{D}$ for all $t \in [t_0, T_{\max})$ and $\mathcal{D}$ is compact, $\sup_{t\in[t_0,T_{\max})} \|z\left(t\right)\| < \infty$. By [7, Chapter 1, Theorem 4.1], the solution extends from $[t_0, T_{\max})$ to $[t_0, \infty)$. Therefore, the solution exists for all $t \geq t_0$ with $z\left(t\right) \in \mathcal{D}$ for all $t \geq t_0$. Consequently, for all $z_0 \in \mathcal{S}$, (3-27) holds for every $t \geq t_0$.

As $t \to \infty$, the bound converges to $\|z\left(t\right)\| \leq \sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}$, i.e., the solution is ultimately bounded with ultimate bound $\mathcal{U}$ by (3-20). Next, recall that $k_{\min} > \lambda_V + \frac{1}{k_2}\rho\left(\sqrt{\frac{\delta}{\lambda_V}}\left(\sqrt{\frac{\lambda_\varphi}{\lambda_1}} + \frac{\lambda_\varphi}{\lambda_1}\right)\right)$. Using the definition of $\overline{\rho}$ and the fact that $\overline{\rho}$ is invertible yields $\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}} < \sqrt{\frac{\lambda_1}{\lambda_\varphi}}\overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \varrho\left(0\right)\right) - \sqrt{\frac{\delta}{\lambda_V}}$, i.e., $\mathcal{U} \subset \mathcal{S}$. Furthermore, since $\frac{\lambda_1}{\lambda_\varphi} \leq 1$ and $\sqrt{\frac{\delta}{\lambda_V}} > 0$, strict monotonicity of ρ gives

$$k_{\min} > \lambda_V + \frac{1}{k_2}\rho\left(\sqrt{\frac{\delta}{\lambda_V}}\left(\sqrt{\frac{\lambda_\varphi}{\lambda_1}} + \frac{\lambda_\varphi}{\lambda_1}\right)\right) > \lambda_V + \frac{1}{k_2}\rho\left(\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}\right),$$

which implies $k_2\left(k_{\min} - \lambda_V\right) - \rho\left(0\right) > \overline{\rho}\left(\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}\right) > 0$. Hence, $\sqrt{\frac{\lambda_1}{\lambda_\varphi}}\overline{\rho}^{-1}\left(k_2\left(k_{\min} - \lambda_V\right) - \rho\left(0\right)\right) - \sqrt{\frac{\delta}{\lambda_V}} > 0$, so $\mathcal{S}$ has strictly positive radius i.e., $\mathcal{S}$ is nonempty. Thus, $\mathcal{U} \subset \mathcal{S} \subset \mathcal{D}$.

Since $z\left(t\right) \in \mathcal{D}$ for all $t \geq t_0$, the images of $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{D}$ under $z \mapsto \kappa$ lie in Ω by (3-18). Consequently, $\kappa\left(t\right) \in \Omega$ for all $t \geq t_0$, and the universal approximation property applied in (3-6) holds for all time.

Since (3-27) holds with constants $\lambda_1, \lambda_\varphi, \lambda_V, \delta$ chosen independently of t_0 and z_0, the right-hand side depends on time only through $(t - t_0)$. Hence, the exponential convergence to $\mathcal{U}$ is uniform (e.g., [76, Definition 1]).

Additionally, since $z(t) \in \mathcal{D}$ for all $t \geq t_0$, $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$. Since $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$, $e, r \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^n\right)$ and $\widetilde{\Theta} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^p\right)$. Therefore, since $q_d, \dot{q}_d \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq 0}; \mathbb{R}^n\right)$ by Assumption 3.1, using (3-3) and (3-4) yields that $q, \dot{q} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq 0}; \mathbb{R}^n\right)$, implying that $g^+(q, \dot{q}, t) \in \mathcal{L}_\infty\left(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0}, \mathbb{R}^m\right)$ by Assumption 3.3. Following (3-23) and the fact that $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$ yields that $\kappa \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{4n}\right)$. Due to the projection operator, $\widehat{\Theta} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^p\right)$. Since $\left(\kappa, \widehat{\Theta}\right) \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq 0}; \mathbb{R}^{4n} \times \mathbb{R}^p\right)$, $\Psi\left(\kappa(t), \widehat{\Theta}(t)\right)$ is bounded. Thus, by (3-7), u is bounded. □

3.1.3 Experiment

Experimental validation was performed on a Freefly Astro quadrotor equipped with a PX4 flight controller at the University of Florida's Autonomy Park outdoor facility. State estimation utilized the onboard EKF2 fusing GPS, optical flow, and Lidar data. A cascaded control architecture was employed, where the proposed ResNet-based controller, implemented as a 50 Hz ROS2 outer loop, generated acceleration commands sent via MAVROS to the PX4's inner-loop controller operating at 400 Hz. The quadrotor autonomously tracked a 15 m × 5 m figure-eight trajectory at 2.5 m altitude for 360 s. The proposed controller was compared against three

Table 3-1. Parameters used in the comparative experiment of SNN, DNN, and ResNet -based adaptive controllers.

	SNN	DNN	ResNet
Neurons	8	2	2
Layers	1	32	2
Blocks	0	0	4
Parameters	254		226
Outer Activation	tanh		
Inner Activation	N/A	Swish	
Shortcut Activation	N/A		Swish
Learning Rate	$\Gamma = 0.05$	$\Gamma = 0.1$	$\Gamma = 0.025$
Search Space Bound	$\overline{\Theta} = 4$	$\overline{\Theta} = 8$	$\overline{\Theta} = 1$
Control Gains	$k_1 = 0.77$, $k_2 = 0.66$, $k_3 = 1\text{e}^{-6}$		

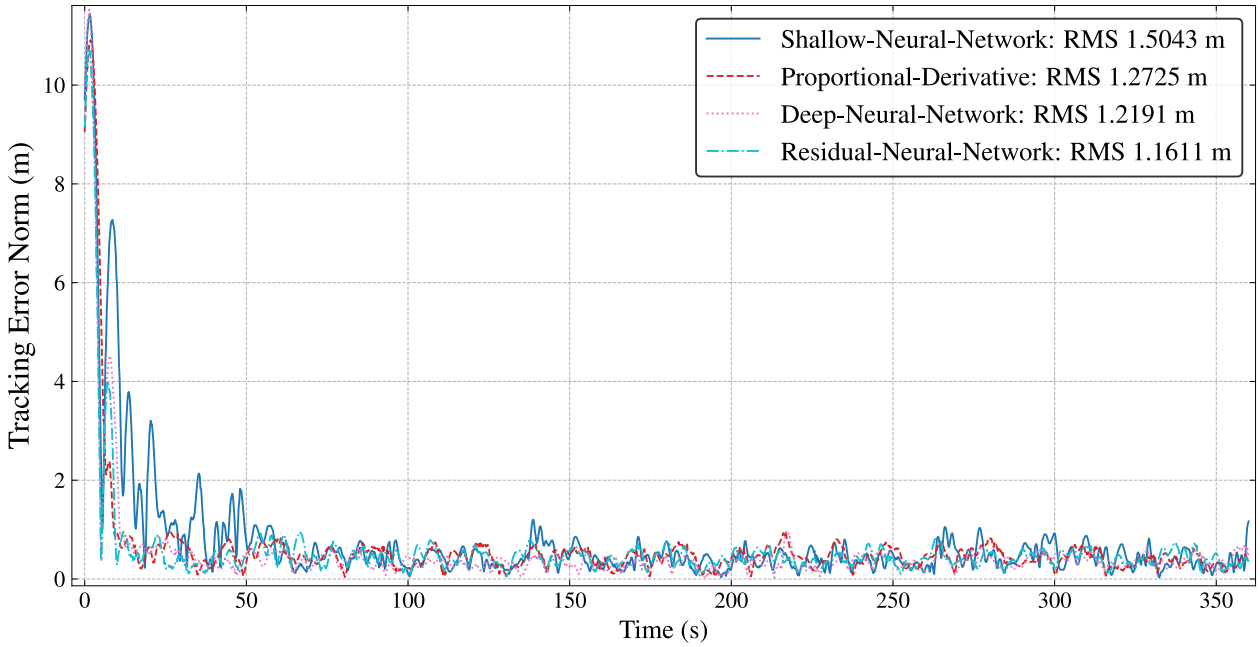

Figure 3-1. Tracking error comparison over the 360-second experiment.

benchmarks: a proportional-derivative (PD) controller, a shallow NN-based adaptive controller (SNN) and a DNN-based adaptive controller (employing Φ instead of Ψ in (3-7)). Controller parameters are provided in Table 3-1. Figure 3-1 illustrates that the proposed ResNet controller reduced tracking error by 22.81%, 8.75%, and 4.76% relative to the SNN, PD, and DNN controllers, respectively, while using approximately 11% fewer parameters compared to the SNN and DNN controllers.

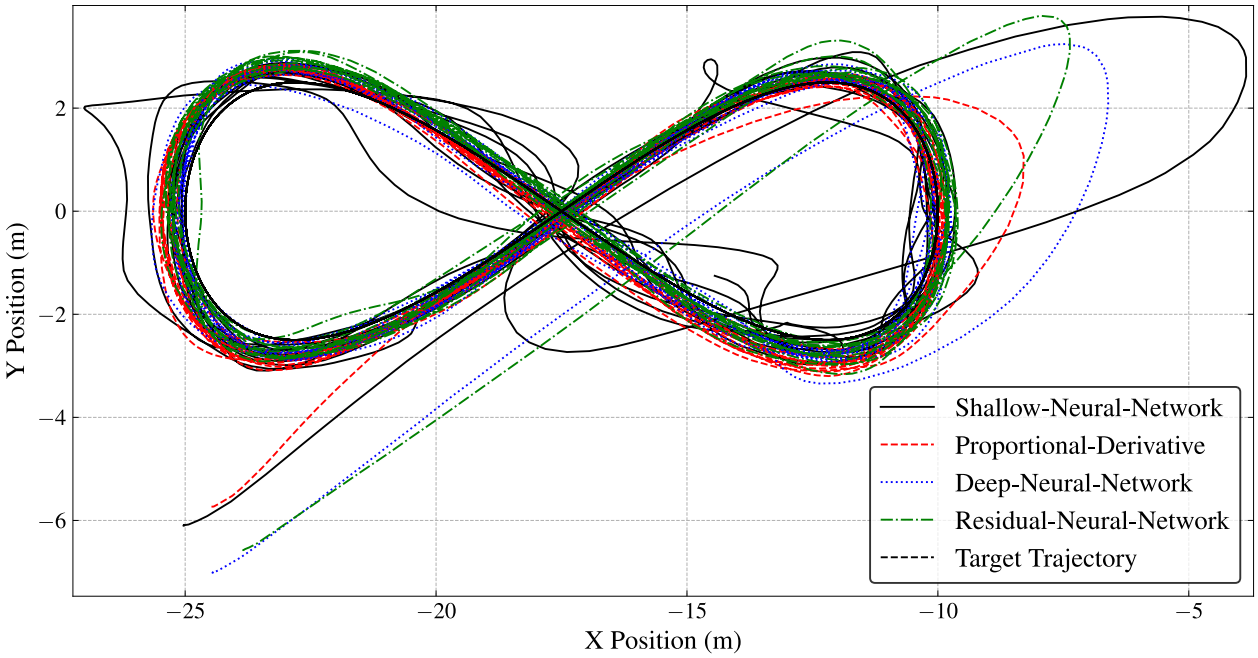

Figure 3-2. Trajectory comparison over the 360-second experiment.

3.2 Second-Order Heterogeneous Multi-Agent Target Tracking without Relative Velocities

Second-order multi-agent target tracking underpins diverse systems, from manipulators to aircraft, and has progressed rapidly in distributed guidance, navigation, and control. Practical deployments rarely provide reliable relative velocity measurements, making full-state feedback assumptions unrealistic and motivating strategies that use only relative positions and onboard sensing. Prior leader–follower and observer-based designs imposed first-order models, required measured relative velocities, restricted agents to linear dynamics, or assumed identical drift dynamics, limiting robustness to heterogeneity and nonlinearity. This work develops a distributed observer for networks with unknown, heterogeneous second-order nonlinear dynamics that operates without relative velocity measurements by introducing an auxiliary distributed filter driven solely by relative positions to estimate the velocities of both target and agents. A Lyapunov analysis establishes exponential regulation of all agents to a neighborhood of the target. Comparative simulations on a six-agent network against an established baseline demonstrate a 67% reduction in RMS tracking error, indicating improved performance under realistic sensing constraints.

3.2.1 Problem Formulation

Consider a network composed of N agents indexed by $\mathcal{V}$, with a static, connected, and undirected communication graph modeled by $\mathcal{G} \triangleq (\mathcal{V}, \mathcal{E})$, where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the edge set. The model for agent $i \in \mathcal{V}$ is given by

$$\ddot{q}_i = f_i\left(q_i, \dot{q}_i, t\right) + u_i, \tag{3-28}$$

where $q_i, \dot{q}_i, \ddot{q}_i \in \mathbb{R}^n$ denote the generalized position, velocity, and acceleration, respectively, $u_i \in \mathbb{R}^n$ denotes the control input, and $f_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ denotes an unknown locally Lipschitz function in q_i and $\dot{q}_i$, uniformly bounded in t. Consider a target with the stable model

$$\ddot{q}_0 = g\left(q_0, \dot{q}_0, t\right), \tag{3-29}$$

such that $\|q_0\| \leq \bar{q}_0 \in \mathbb{R}_{\geq 0}$ and $\|\dot{q}_0\| \leq \bar{\dot{q}}_0 \in \mathbb{R}_{\geq 0}$, where $q_0, \dot{q}_0, \ddot{q}_0 \in \mathbb{R}^n$ denote the generalized position, velocity, and acceleration, respectively, and $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ denotes an unknown locally Lipschitz function in q_0 and $\dot{q}_0$, uniformly bounded in t. Each agent is capable of measuring only the relative position between itself and its neighbors, which is given by

$$d_{ij} \triangleq q_j - q_i, \quad \forall j \in \mathcal{N}_i. \tag{3-30}$$

The objective is to develop a distributed controller capable of regulating each agent to the target, based on only the relative position measurement. To quantify the target regulation objective, let the tracking error $e_i \in \mathbb{R}^n$ of agent $i \in \mathcal{V}$ be

$$e_i \triangleq q_0 - q_i. \tag{3-31}$$

Based on the subsequent stability analysis, we also define the auxiliary tracking error $r_i \in \mathbb{R}^n$ of agent $i \in \mathcal{V}$ as

$$r_i \triangleq \dot{e}_i + k_1 e_i, \tag{3-32}$$

where $k_1 \in \mathbb{R}_{>0}$ is a user-defined gain. The following section provides the control design to achieve the stated objective.

3.2.2 Control Synthesis

Since at least one agent must be able to measure e_i, we let $b_i \in \{0, 1\}$ denote the indicator of whether an agent $i \in \mathcal{V}$ is capable of measuring e_i. Let the relative position $\eta_i \in \mathbb{R}^n$ and relative velocity $\zeta_i \in \mathbb{R}^n$ be defined as

$$\eta_i \triangleq \sum_{j \in \mathcal{N}_i} d_{ij} + b_i e_i, \tag{3-33}$$

and

$$\zeta_i \triangleq \dot{\eta}_i = \sum_{j \in \mathcal{N}_i} \dot{d}_{ij} + b_i \dot{e}_i, \tag{3-34}$$

respectively, for all $i \in \mathcal{V}$.

Define the interaction matrix $\mathcal{H} \triangleq \left(\mathcal{L}_{\mathcal{G}} + \text{diag}\{b_1, \ldots, b_N\}\right) \otimes I_n \in \mathbb{R}^{nN \times nN}$. Using (3-33) and the definition of the graph Laplacian and interaction matrix, the relative positions can be expressed in an ensemble form as

$$\eta = \mathcal{H} e, \tag{3-35}$$

where $\eta \triangleq [\eta_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$ and $e \triangleq [e_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Based on (3-35), the relative velocities can be expressed in an ensemble form as

$$\zeta = \mathcal{H} \dot{e}, \tag{3-36}$$

where $\zeta \triangleq [\zeta_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$.

Remark 3.3. Since the communication graph is connected and $b_i \neq 0$ for at least one i, $\mathcal{H}$ is a positive-definite M-matrix. [80, Corollary 4.33].

3.2.2.1 Distributed Observer Development

Since the relative velocities are unknown, we develop a distributed observer that generates estimates of η_i and ζ_i, given by $\hat{\eta}_i, \hat{\zeta}_i \in \mathbb{R}^n$, respectively. The corresponding relative position estimation error $\tilde{\eta}_i \in \mathbb{R}^n$ and the relative velocity estimation error $\tilde{\zeta}_i \in \mathbb{R}^n$ are quantified as

$$\tilde{\eta}_i \triangleq \hat{\eta}_i - \eta_i, \quad \tilde{\zeta}_i \triangleq \hat{\zeta}_i - \zeta_i, \tag{3-37}$$

respectively, for all $i \in \mathcal{V}$. To facilitate the subsequent development, we introduce the auxiliary estimation error $\tilde{r}_i \in \mathbb{R}^n$ defined as

$$\tilde{r}_i \triangleq \dot{\tilde{\eta}}_i + k_3 \tilde{\eta}_i + \rho_i, \tag{3-38}$$

for all $i \in \mathcal{V}$, where $k_3 \in \mathbb{R}_{>0}$ is a user-defined gain, and $\rho_i \in \mathbb{R}^n$ denotes the output from a dynamic filter which compensates for the fact that the state ζ_i is not measurable. Based on the subsequent stability analysis, ρ_i is designed as the output of the dynamic filter

$$\dot{\rho}_i \triangleq \tilde{\eta}_i - (k_3 + k_4)\, \tilde{r}_i - k_5 \rho_i, \quad \rho_i\,(t_0) = 0, \tag{3-39}$$

where $k_4, k_5 \in \mathbb{R}_{>0}$ are user-defined gains. Substituting (3-38) into (3-39) and integrating the resulting equation from t_0 to t, (3-39) is implemented as

$$\rho_i(t) = \left(1 - k_3^2 - k_3 k_4\right) \int_{t_0}^{t} \tilde{\eta}_i\,(\tau)\, \mathrm{d}\tau - (k_3 + k_4 + k_5) \int_{t_0}^{t} \rho_i\,(\tau)\, \mathrm{d}\tau - (k_3 + k_4)\, \tilde{\eta}_i\,(t)\,, \tag{3-40}$$

for all $i \in \mathcal{V}$. Based on the subsequent analysis, we design the distributed observer as

$$\begin{aligned} \dot{\hat{\eta}}_i &= \hat{\zeta}_i, \quad \hat{\eta}_i\,(t_0) = 0, \\ \dot{\hat{\zeta}}_i &= \sum_{j \in \mathcal{N}_i} \left(u_j - u_i\right) - b_i u_i + (2k_3 + k_4 + k_5)\, \rho_i - \left(2 - k_3^2\right) \tilde{\eta}_i, \quad \hat{\zeta}_i\,(t_0) = 0, \end{aligned} \tag{3-41}$$

for all $i \in \mathcal{V}$. Using the definitions of the graph Laplacian and interaction matrix, (3-41) can be expressed as

$$\begin{aligned} \dot{\hat{\eta}} &= \hat{\zeta}, \\ \dot{\hat{\zeta}} &= -\mathcal{H}u + (2k_3 + k_4 + k_5)\,\rho - \left(2 - k_3^2\right)\tilde{\eta}, \end{aligned} \tag{3-42}$$

where $u \triangleq [u_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $\hat{\eta} \triangleq [\hat{\eta}_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $\hat{\zeta} \triangleq \left[\hat{\zeta}_i\right]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $\rho \triangleq [\rho_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, and $\tilde{\eta} \triangleq [\tilde{\eta}_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$. Substituting (3-36)-(3-39) and (3-42) into the time-derivative of (3-38) yields

$$\dot{\tilde{r}} = (k_3 + k_4)\,\rho - \tilde{\eta} - \mathcal{H}\left(G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right)\right) - k_4\tilde{r}, \tag{3-43}$$

where $G\left(q_0, \dot{q}_0, t\right) \triangleq \left(1_N \otimes g\left(q_0, \dot{q}_0, t\right)\right) \in \mathbb{R}^{nN}$, $F\left(q, \dot{q}, t\right) \triangleq \left[f_i\left(q_i, \dot{q}_i, t\right)\right]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $q \triangleq [q_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, and $\tilde{r} \triangleq [\tilde{r}_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$.

3.2.2.2 Distributed Controller Development

Based on the subsequent analysis, we design the distributed controller as

$$u_i = k_1 k_2 \eta_i + k_2\hat{\zeta}_i + k_2 k_3\tilde{\eta}_i + k_2\rho_i, \tag{3-44}$$

where $k_2 \in \mathbb{R}_{>0}$ is a user-defined gain. Substituting (3-35), (3-36), and (3-38) into (3-44) yields

$$u = k_2\mathcal{H}r + k_2\tilde{r}. \tag{3-45}$$

Substituting (3-28), (3-29), (3-31), (3-32), and (3-45) into the time-derivative of (3-32) yields

$$\dot{r} = G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right) - k_2\mathcal{H}r - k_2\tilde{r} + k_1 r - k_1^2 e. \tag{3-46}$$

The block diagrammatic representation for a combined implementation of the distributed filter, observer, and controller is shown in Figure 3-3. The following section provides the stability

analysis for the distributed filter, observer, and controller. To facilitate the stability analysis, we introduce the following lemma.

Lemma 3.1. *Let $L \in \mathbb{R}_{\geq 0}$ be defined as $L \triangleq \max\{L_1, \ldots, L_N\}$, where $L_i \in \mathbb{R}_{\geq 0}$ is the drift Lipschitz constant for agent $i \in \mathcal{V}$ defined for $q, Q_0 \in \mathcal{D}_1 \subseteq \mathbb{R}^{nN}$ and $\dot{q}, \dot{Q}_0 \in \mathcal{D}_2 \subseteq \mathbb{R}^{nN}$, where $Q_0 \triangleq 1_N \otimes q_0 \in \mathbb{R}^{nN}$. There exists some constant $K \in \mathbb{R}_{\geq 0}$ such that*

$$\|G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t)\| \leq L\|r\| + L(1 + k_1)\|e\| + NK. \tag{3-47}$$

Proof. Adding and subtracting $F(q_0, \dot{q}_0, t)$ to $G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t)$ and using the triangle inequality yields

$$\begin{aligned} \|G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t)\| \leq &\left\|F\left(Q_0, \dot{Q}_0, t\right) - F(q, \dot{q}, t)\right\|, \\ &+ \left\|G(q_0, \dot{q}_0, t) - F\left(Q_0, \dot{Q}_0, t\right)\right\|. \end{aligned} \tag{3-48}$$

Since g and f are continuous functions on $\mathcal{D}_1$ and $\mathcal{D}_2$, there exists some constant $K \in \mathbb{R}_{\geq 0}$ such that

$$\max_{i \in \mathcal{V}} \|g(q_0, \dot{q}_0, t) - f_i(q_0, \dot{q}_0, t)\| \leq K, \tag{3-49}$$

for all $t \in \mathbb{R}^n$. Applying (3-49) and the Lipschitz property of f_i, $\forall i \in \mathcal{V}$ to (3-48), and using the definitions of e and $\dot{e}$ yields

$$\|G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t)\| \leq L(\|e\| + \|\dot{e}\|) + NK. \tag{3-50}$$

Substituting (3-32) into (3-50) and bounding the resulting inequality yields (3-47). □

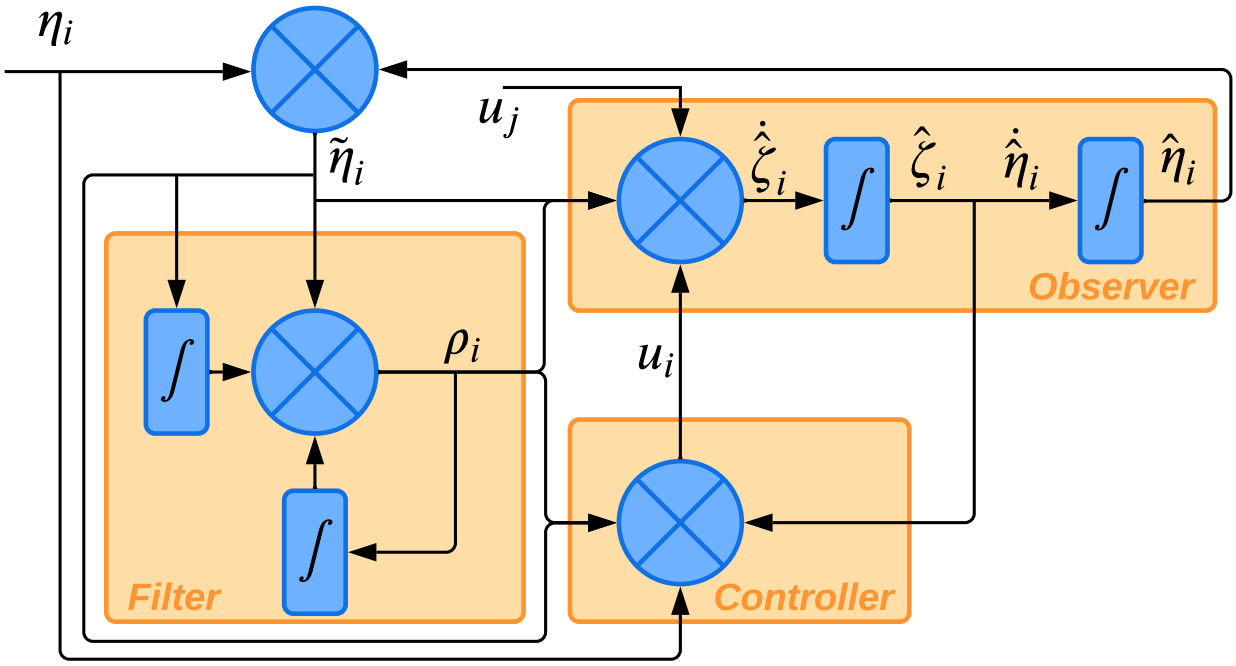

Figure 3-3. Block diagram of algorithm implementation for agent $i \in \mathcal{V}$. Control gains have been omitted for clarity.

3.2.3 Stability Analysis

Define the state vector $\xi : \mathbb{R}_{\geq 0} \to \mathbb{R}^{5nN}$ as $\xi \triangleq \begin{bmatrix} e^\top & r^\top & \tilde{\eta}^\top & \tilde{r}^\top & \rho^\top \end{bmatrix}^\top$. Using (3-32), (3-38), (3-39), (3-43), and (3-46) yields the closed-loop error system

$$\dot{\xi} = \begin{bmatrix} r - k_1 e \\ G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right) - k_2 \mathcal{H} r - k_2 \tilde{r} + k_1 r - k_1^2 e \\ \tilde{r} - k_3 \tilde{\eta} - \rho \\ \left(k_3 + k_4\right) \rho - \tilde{\eta} - \mathcal{H}\left(G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right)\right) - k_4 \tilde{r} \\ \tilde{\eta} - \left(k_3 + k_4\right) \tilde{r} - k_5 \rho \end{bmatrix}. \tag{3-51}$$

Consider the continuously differentiable Lyapunov function candidate $V : \mathbb{R}^{5nN} \to \mathbb{R}_{\geq 0}$ defined as

$$V\left(\xi\right) \triangleq \frac{1}{2} \xi^\top \xi. \tag{3-52}$$

Let $\gamma, \delta \in \mathbb{R}_{>0}$ be defined as $\gamma \triangleq \lambda_{\max}\left\{\mathcal{L}_{\mathcal{G}}\right\} + 1$ and $\delta \triangleq \frac{NK^2(1+\gamma)}{2}$, respectively.

Since the local Lipschitz property of F holds over a subset of $\mathbb{R}^{nN}$, the subsequent stability analysis requires ensuring $q \in \mathcal{D}_1$ and $\dot{q} \in \mathcal{D}_2$, $\forall i \in \mathcal{V}, \forall t \in \mathbb{R}_{\geq 0}$. This is achieved by yielding a local stability result which constrains the states as $\|\xi\| \leq \chi$, where $\chi \in \mathbb{R}_{>0}$. Based on the subsequent stability analysis, we define $\mathcal{S} \triangleq \left\{ \varrho : \|\varrho\| \leq \sqrt{\chi^2 - \frac{\delta}{k}} \right\}$, where $k \in \mathbb{R}_{>0}$.

Theorem 3.2. *Consider the dynamical systems described by* (3-28) *and* (3-29)*. For any initial conditions of the states* $\|\xi(t_0)\| \in \mathcal{S}$*, the filter given by* (3-40)*, observer given by* (3-42)*, and controller given by* (3-44) *guarantee that there exists a constant* k *such that*

$$\|\xi(t)\| \leq \sqrt{\left(\|\xi(t_0)\|^2 - \frac{\delta}{k} \right) \mathrm{e}^{-2k(t-t_0)} + \frac{\delta}{k}}, \tag{3-53}$$

$\forall t \in [t_0, \infty)$*, provided that the control gains* k_1, k_2*, and* k_4 *are selected such that* $k_1 > 1$*,* $-\kappa_1 \left(1 - k_1^2\right) - L(1 + k_1)(1 + \gamma) + 2k_1 > 0$*,* $k_2 > \frac{-k_1^2 + \kappa_1 (k_1(2+L) + L(3+\gamma) + N) + 1}{\kappa_1 (2\lambda_{\min}\{\mathcal{H}\} - \kappa_2)}$*, and* $k_4 > \frac{k_2 + \gamma\kappa_2 (L(2+k_1) + N)}{2\kappa_2}$*, where* $\kappa_1 > 0$ *and* $\kappa_2 > 2\lambda_{\min}\{\mathcal{H}\}$ *are constants.*

Proof. Substituting (3-51) into the time-derivative of (3-52) yields

$$\begin{aligned} \dot{V}(\xi) = &-k_1 e^\top e - k_2 r^\top \mathcal{H} r - k_3 \tilde{\eta}^\top \tilde{\eta} - k_4 \tilde{r}^\top \tilde{r} - k_5 \rho^\top \rho + k_1 r^\top r + \left(1 - k_1^2\right) e^\top r \\ &- k_2 r^\top \tilde{r} - \tilde{r}^\top \left(\mathcal{H} \left(G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t) \right) \right) + r^\top \left(G(q_0, \dot{q}_0, t) - F(q, \dot{q}, t) \right). \end{aligned} \tag{3-54}$$

Applying Young's inequality to $\left(1 - k_1^2\right) e^\top r$ yields

$$\left(1 - k_1^2\right) e^\top r \leq \frac{\kappa_1 \left|1 - k_1^2\right|}{2} \|e\|^2 + \frac{\left|1 - k_1^2\right|}{2\kappa_1} \|r\|^2, \tag{3-55}$$

where $\kappa_1 \in \mathbb{R}_{>0}$. Similarly, we have that $-k_2 r^\top \tilde{r}$ may be upper bounded as

$$-k_2 r^\top \tilde{r} \leq \frac{k_2 \kappa_2}{2} \|r\|^2 + \frac{k_2}{2\kappa_2} \|\tilde{r}\|^2, \tag{3-56}$$

where $\kappa_2 \in \mathbb{R}_{>0}$. We can upper bound $r^\top \left(G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right)\right)$ by applying Lemma 3.1 as

$$r^\top \left(G\left(q_0, \dot{q}_0, t\right) - F\left(q, \dot{q}, t\right)\right) \leq \frac{L\left(1 + k_1\right)}{2} \|e\|^2 + \frac{L\left(3 + k_1\right) + N}{2} \|r\|^2 + \frac{NK^2}{2}. \tag{3-57}$$

Based on the definition of the graph Laplacian, we can upper-bound the norm of $\mathcal{H}$ as

$$\|\mathcal{H}\| \leq \gamma. \tag{3-58}$$

Similarly, we can upper bound $\tilde{r}^\top \left(\mathcal{H}\left(G(q_0, \dot{q}_0, t) - F\left(q, \dot{q}, t\right)\right)\right)$ by using (3-58), applying Lemma 3.1, and using Young's inequality, yielding

$$\begin{aligned} \tilde{r}^\top \left(\mathcal{H}\left(G(q_0, \dot{q}_0, t) - F\left(q, \dot{q}, t\right)\right)\right) \leq \gamma \left(\frac{L}{2} \|r\|^2 + \frac{L\left(1 + k_1\right)}{2} \|e\|^2\right) \\ + \gamma \left(\frac{L\left(2 + k_1\right) + N}{2} \|\tilde{r}\|^2 + \frac{NK^2}{2}\right) \end{aligned} \tag{3-59}$$

Thus, we may upper bound (3-54) as

$$\dot{V}\left(\xi\right) \leq -k \|\xi\|^2 + \frac{NK^2\left(1 + \gamma\right)}{2}, \tag{3-60}$$

by using (3-55)-(3-57), (3-59), and invoking the gain conditions of k_1, k_2, and k_4. Using (3-52), we can upper bound (3-60) as

$$\dot{V}\left(\xi\right) \leq -2kV\left(\xi\right) + \delta. \tag{3-61}$$

Solving the differential inequality in (3-61) yields (3-53).

Further upper bounding the right side of (3-53) yields $\|\xi\left(t\right)\| \leq \sqrt{\|\xi\left(t_0\right)\|^2 + \frac{\delta}{k}}$ for all $t \in \mathbb{R}_{\geq 0}$. Consider the set $\bar{\mathcal{B}}_\chi \triangleq \{\iota : \|\iota\| \leq \chi\}$. If $\xi\left(t_0\right) \in \mathcal{S}$, then $\sqrt{\|\xi\left(t_0\right)\|^2 + \frac{\delta}{k}} \leq \chi$ which implies that $\xi \in \bar{\mathcal{B}}_\chi$ always holds. Since $\|\xi\left(t\right)\| \leq \chi$ implies $\|e\|, \|r\| \leq \chi$, the relations $\|q\| \leq \chi + N\bar{q}_0$ and $\|\dot{q}\| \leq \left(1 + k_1\right)\chi + N\bar{\dot{q}}_0$ hold. Thus, selecting χ such that

$\mathcal{D}_1 = \{\iota_1 : \iota_1 \leq \chi + N\bar{q}_0\}$ and $\mathcal{D}_2 = \left\{\iota_2 : \iota_2 \leq (1 + k_1)\,\chi + N\bar{\dot{q}}_0\right\}$ ensures that $q \in \mathcal{D}_1$ and $\dot{q} \in \mathcal{D}_2$ for all $t \in \mathbb{R}_{\geq 0}$, thus always satisfying the local Lipschitz property. If F is globally Lipschitz, then the stability result holds for all initial conditions.

Similarly, since $\|\rho\|, \|\tilde{\eta}\|, \|\tilde{r}\| \leq \chi$, $\rho, \tilde{\eta}, \tilde{r} \in \mathcal{L}_\infty$. Thus, using (3-38) yields $\tilde{\zeta} \in \mathcal{L}_\infty$. Since $e, r \in \mathcal{L}_\infty$, using (3-32) yields $\dot{e} \in \mathcal{L}_\infty$. Since $e, \dot{e} \in \mathcal{L}_\infty$, using (3-35) and (3-36) yields $\eta, \zeta \in \mathcal{L}_\infty$. Since $\eta, \tilde{\eta}, \zeta, \tilde{\zeta} \in \mathcal{L}_\infty$, using (3-37) yields $\hat{\eta}, \hat{\zeta} \in \mathcal{L}_\infty$. Since $\eta_i, \hat{\zeta}_i, \tilde{\eta}_i, \rho_i \in \mathcal{L}_\infty$, using (3-44) yields $u \in \mathcal{L}_\infty$. Since $u, \rho, \tilde{\eta}, \hat{\zeta} \in \mathcal{L}_\infty$, using (3-42) yields $\dot{\hat{\zeta}}, \dot{\hat{\eta}} \in \mathcal{L}_\infty$. Therefore, all implemented signals are bounded. □

3.2.4 Simulation

The performance of the developed controller is demonstrated through comparative simulations, with the method in [9, eq.40] as the baseline. The controller and observer in the baseline method are given by

$$\begin{aligned} u_i &= \sigma\eta_i + \sigma\hat{\zeta}_i, \\ \dot{\hat{\eta}}_i &= \hat{\zeta}_i - \ell\tilde{\eta}_i, \\ \dot{\hat{\zeta}}_i &= \sum_{j\in\mathcal{N}_i} \left(u_j - u_i\right) - b_i u_i - \ell^2\tilde{\eta}_i, \end{aligned} \tag{3-62}$$

where $\sigma, \ell \in \mathbb{R}_{>0}$ are control and observer gains, respectively. The baseline method was developed assuming the agent dynamics are homogeneous and identical to the target, i.e., $f_i = g$ for all agents $i \in \mathcal{V}$. The developed method does not rely on this assumption, and thus also accounts for heterogeneity in the dynamics. To demonstrate the efficacy of the developed method, six simulations are performed, with each simulation having a different degree of heterogeneity. Each simulation is performed for 60 seconds, where six agents are tasked with tracking a moving target. The drift dynamics for agent $i \in \mathcal{V}$ in (3-28) are defined as those of a damped and driven harmonic oscillator, given by [9, eq.40]

$$f_i\left(q_i, \dot{q}_i, t\right) = -a_i \sin\left(q_i\right) - b_i\dot{q}_i + c_i \cos\left(d_i t\right). \tag{3-63}$$

To show how varying levels of heterogeneity impact the results, we choose the parameters a_i, b_i, c_i, d_i from the uniform distribution $U(0, \omega)$. In each simulation, we use different values of ω to create these ranges, i.e., $\omega \in \{0, 2.5, 5, 7.5, 10, 1000\}$, where a broader range of these parameters (i.e., higher ω) suggests a more diverse set of dynamics that controlled agents might encounter. Additionally, all simulations were conducted with an additive measurement noise sampled from a Gaussian distribution with variance 0.5.

The initial conditions of the agents are selected so that they form an equilateral hexagon of radius 0.5, centered at $(x, y, z) = (0.5, 0.5, 0)$. The dynamics for the target are also given by (3-63), with initial conditions

$$x_0(0) = y_0(0) = z_0(0) = 0.5,$$
$$\dot{x}_0(0) = \dot{y}_0(0) = 0, \quad \dot{z}_0(0) = 0.1.$$

Three agents are selected so that $b_i = 1$, with the remaining three agents having $b_i = 0$, i.e. only three agents could measure the relative positions between themselves and the target. This configuration results in the indicator matrix being given by diag $\{0, 1, 0, 1, 0, 1\}$. The graph Laplacian for the network is given by

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} 3 & 0 & -1 & -1 & -1 & 0 \\ 0 & 3 & 0 & -1 & -1 & -1 \\ -1 & 0 & 3 & 0 & -1 & -1 \\ -1 & -1 & 0 & 3 & 0 & -1 \\ -1 & -1 & -1 & 0 & 3 & 0 \\ 0 & -1 & -1 & -1 & 0 & 3 \end{bmatrix}.$$

An adaptive Monte-Carlo method is used to select the final gains in each simulation. A total of 25,000 trials are run for each value of ω, with the gains being selected from the uniform

distribution $U\,(0.001, 50)$. The cost function is defined as

$$J = \int_0^{t_f} \left(\|e\,(t)\|^2 + \tau^2 \left\|\tilde{\zeta}\,(t)\right\|^2 \right) \mathrm{d}t,$$

where $\tau = 0.25$ is a scaling factor.

Every 1000 trials, the top 50% performing trials are selected and the corresponding gains are used to determine the new set for random gain initialization. To ensure a fair comparison, both scenarios employ the same adaptive Monte Carlo simulation approach, and all the initial conditions, dynamics, and simulation parameters for both agents and targets are kept identical. The simulations are ran in parallel to ensure that the random dynamic parameters are identical for both simulations. The final gains selected for the both the developed controller and the baseline controller are shown in Table 3-2.

Considering the limitations of the baseline method outlined in [9, eq.40] for heterogeneous scenarios, we conducted a homogeneous comparison ($\omega = 0$), the results of which are depicted in Figure 3-4. These results validate the efficacy of our method even in the absence of heterogeneity. To further test our methods effectiveness, we increased ω up to 1000. As shown in Table 3-2, our controller maintains satisfactory tracking performance without significant degradation, even under high heterogeneity levels.

Figure 3-5 shows the tracking errors using our developed method at $\omega = 2.5$, with Table 3-3 listing the corresponding randomly initialized parameters. As evident from Figure 3-5, all agents effectively track the target with $\|e_i\| < 0.1$ after approximately 0.5 seconds.

Table 3-2 provides the Root Mean Square (RMS) and maximum tracking and velocity estimation errors for the six simulations using both developed and baseline methods. The baseline method shows a performance decline with increasing dynamics heterogeneity (i.e., higher ω) in contrast to the developed method. Particularly at $\omega = 10$, the developed method significantly outperforms the baseline method, as shown in Figure 3-6, achieving a 67% improvement in the RMS tracking error.

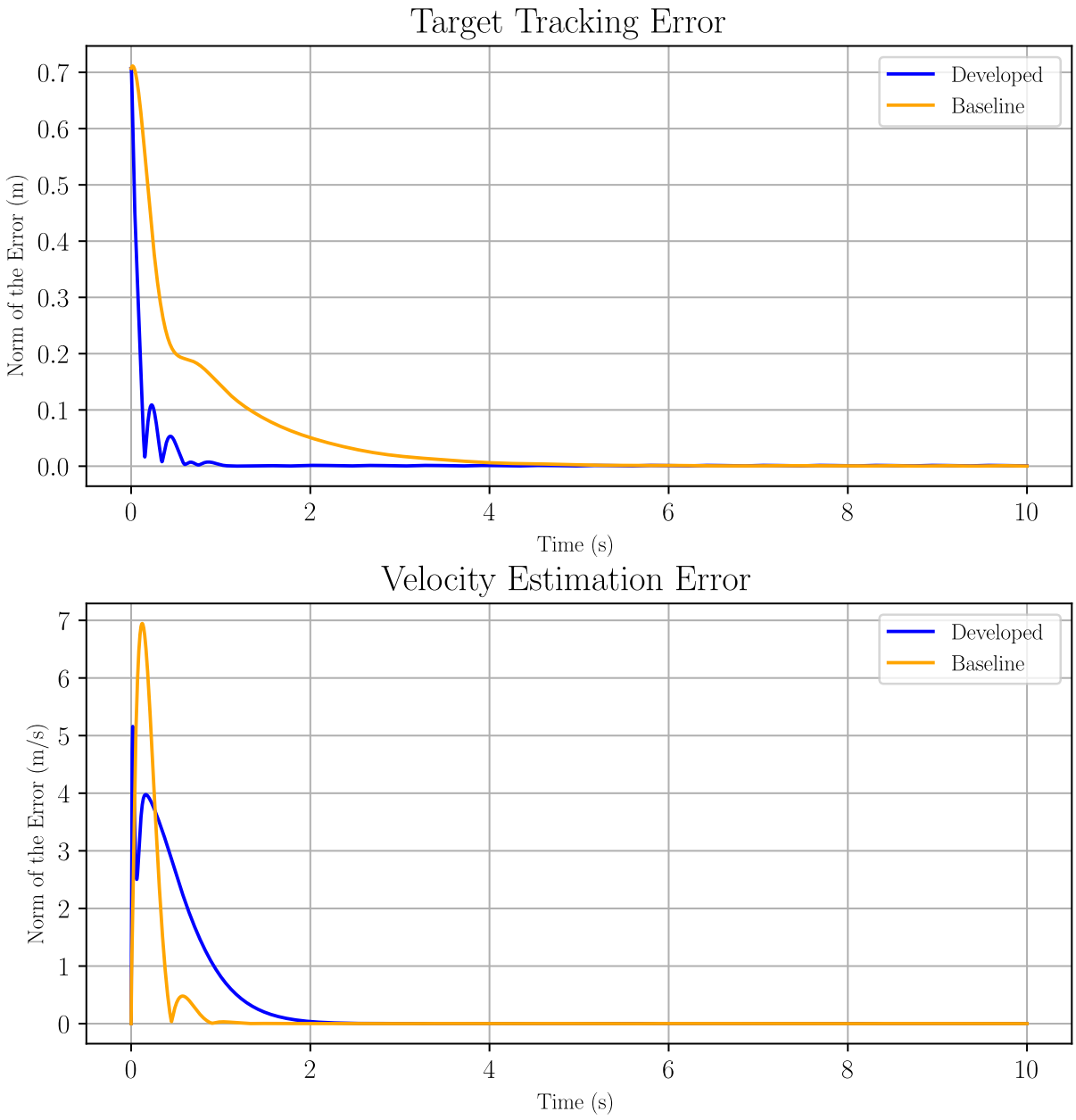

Figure 3-4. Plots of the tracking and velocity estimation errors for the homogeneous comparison ($\omega = 0$).

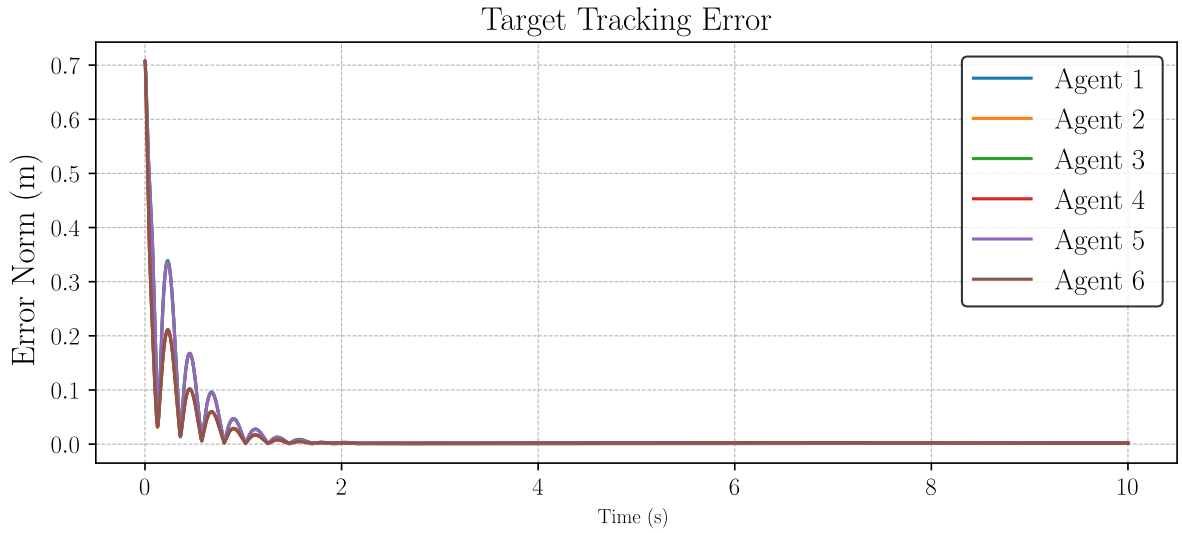

Figure 3-5. Plot of the tracking errors for each agent, when the dynamic parameters a_i, b_i, c_i, d_i are selected from the uniform distribution $U(0, 2.5)$.

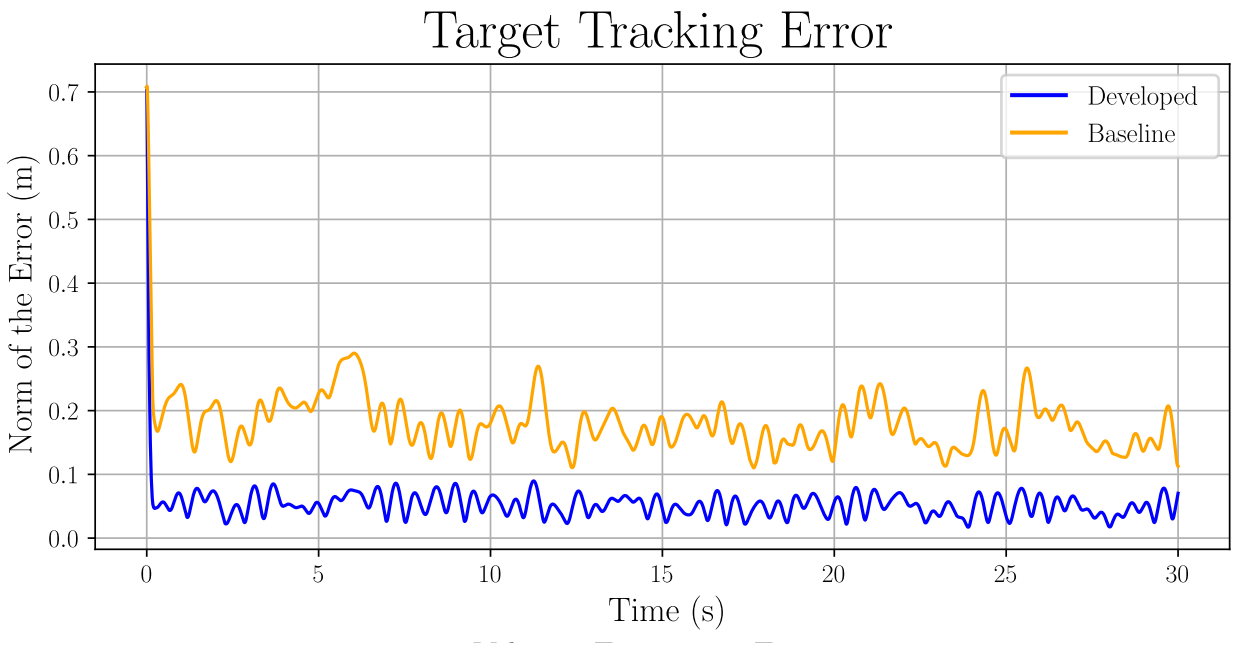

Figure 3-6. Comparative plots of the average tracking errors, when the dynamic parameters a_i, b_i, c_i, d_i are selected from the uniform distribution $U(0, 10)$.

Table 3-2. Result Comparison

ω	Method	Gains	Error	$\|e\|^2$	$\|\tilde{\zeta}\|^2$
0	Developed	39.35, 10.24, 4.01, 1.05, 41.25	RMS	0.04	0.87
			Max	0.70	5.47
	Baseline	36.21, 5.20	RMS	0.11	0.87
			Max	0.71	5.47
2.5	Developed	37.45, 14.34, 3.17, 2.51, 39.29	RMS	0.04	1.32
			Max	0.70	13.37
	Baseline	35.23, 6.91	RMS	0.42	1.50
			Max	0.91	10.32
5.0	Developed	30.62, 22.11, 0.34, 8.94, 31.43	RMS	0.03	1.24
			Max	0.70	17.70
	Baseline	35.11, 13.11	RMS	0.23	1.15
			Max	0.71	20.00
7.5	Developed	18.97, 22.39, 0.13, 16.21, 11.45	RMS	0.04	1.52
			Max	0.70	25.94
	Baseline	34.53, 14.42	RMS	0.18	1.50
			Max	0.71	22.44
10	Developed	9.75, 32.41, 0.05, 20.28, 23.45	RMS	0.06	1.78
			Max	0.70	32.65
	Baseline	34.63, 16.45	RMS	0.18	1.80
			Max	0.70	22.93
1000	Developed	9.75, 32.41, 0.05, 20.28, 23.45	RMS	0.31	14.13
			Max	0.94	37.95
	Baseline	34.63, 16.45	RMS	Unstable	
			Max	Unstable	

Table 3-3. Heterogeneity Parameters for $\omega = 2.5$

Agent	$a = (a_x, a_y, a_z)$	$b = (b_x, b_y, b_z)$
	$c = (c_x, c_y, c_z)$	$d = (d_x, d_y, d_z)$
Agent 1	2.27, 1.33, 1.86	0.91, 1.86 ,1.76
	0.46, 1.96, 0.55	0.46, 1.96, 0.55
Agent 2	0.26, 1.07, 0.87	0.70, 1.27, 0.92
	0.88, 1.36, 0.63	1.14, 0.06, 0.87
Agent 3	1.33, 2.34, 0.88	1.10, 1.42, 1.45
	2.42, 0.95, 2.24	1.00, 0.81, 0.07
Agent 4	0.56, 0.83, 2.40	0.82, 1.44, 1.39
	1.21, 2.36, 0.82	2.35, 1.14, 2.19
Agent 5	1.77, 1.97, 2.40	2.28, 2.25, 0.29
	1.38, 1.52, 0.99	1.82, 2.40, 2.17
Agent 6	2.17, 2.17, 1.36	1.80, 1.34, 0.98
	0.91, 1.51, 1.53	1.52, 1.37, 2.34
Target	0.23, 2.08, 0.88	0.93, 2.35, 1.92
	1.67, 1.76, 0.25	1.82, 1.84, 2.36

CHAPTER 4 DISTRIBUTED RISE-BASED CONTROL FOR EXPONENTIAL HETEROGENEOUS MULTI-AGENT TARGET TRACKING OF SECOND-ORDER NONLINEAR SYSTEMS

Multi-agent target tracking for second-order nonlinear systems faces challenges from network topology, communication limits, unmodeled dynamics, and disturbances, demanding robust methods with stability guarantees. The Robust Integral of the Sign of the Error (RISE) framework achieves exponential tracking error convergence with continuous inputs under time-varying disturbances by integrating a signum term and employing a P-function in Lyapunov analysis, yet extending traditional P-function constructions to multi-agent ensembles is difficult because stability is analyzed at the network level. Earlier multi-agent RISE tracking designs coupled the interaction matrix to the signum term and required two-hop communication, while related applications focused on flocking or stabilization rather than target tracking. Addressing these gaps, a distributed RISE-based controller is formulated that embeds the graph interaction matrix directly in the Lyapunov function, eliminating two-hop communication and motivating a new P-function that incorporates the interaction matrix while preserving key 1-norm identities; a nonsmooth analysis establishes semi-global exponential convergence of all agents to the target despite heterogeneity and disturbances. Core contributions include a distributed design relying only on local information exchange, a multi-agent-tailored P-function construction, and exponential convergence guarantees under uncertainties.

4.1 Problem Formulation

4.1.1 System Dynamics

Consider a multi-agent system composed of N agents indexed by $\mathcal{V}$, and consider a target agent indexed by $\{0\}$. The dynamics for agent $i \in \mathcal{V}$ are given by

$$\ddot{q}_i = f_i\left(q_i, \dot{q}_i, t\right) + g_i\left(q_i, \dot{q}_i, t\right) u_i\left(t\right) + d_i\left(t\right), \tag{4-1}$$

where: $q_i, \dot{q}_i, \ddot{q}_i \in \mathbb{R}^n$ denote the agents' unknown generalized position, velocity, and acceleration, respectively; the unknown functions $f_i : \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^n$ and $d_i : [0, \infty) \to \mathbb{R}^n$ represent

modeling uncertainties and exogenous disturbances, respectively; $g_i : \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^{n \times m_i}$ denotes a known control effectiveness matrix; and $u_i : [0, \infty) \to \mathbb{R}^{m_i}$ denotes the control input. Here, the functions f_i are of class C^3, the mappings $t \mapsto f_i(q_i, \dot{q}_i, t)$, $t \mapsto \nabla f_i(q_i, \dot{q}_i, t)$, and $t \mapsto \nabla^2 f_i(q_i, \dot{q}_i, t)$ are uniformly bounded, the functions g_i have full row-rank, are of class C^3, the mappings $t \mapsto g_i(q_i, \dot{q}_i, t)$, $t \mapsto \nabla g_i(q_i, \dot{q}_i, t)$, and $t \mapsto \nabla^2 g_i(q_i, \dot{q}_i, t)$ are uniformly bounded, and the functions d_i are of class C^2, and there exist known constants $\overline{d}_i, \overline{\dot{d}}_i, \overline{\ddot{d}}_i \in \mathbb{R}_{>0}$ such that $\|d_i(t)\| \leq \overline{d}_i$, $\left\|\dot{d}_i(t)\right\| \leq \overline{\dot{d}}_i$, and $\left\|\ddot{d}_i(t)\right\| \leq \overline{\ddot{d}}_i$ for all $t \in [0, \infty)$, for all $i \in \mathcal{V}$. By the full row-rank property, the existence of the right Moore-Penrose inverse $g_i^+ : \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^{m_i \times n}$ is ensured, for all $i \in \mathcal{V}$, where the functions g_i^+ are of class C^1, and the mappings $t \mapsto g_i^+(q_i, \dot{q}_i, t)$ and $t \mapsto \nabla g_i^+(q_i, \dot{q}_i, t)$ are uniformly bounded for all $i \in \mathcal{V}$.

Remark 4.1. For addressing an uncertain g_i, potential extensions might leverage techniques found in [95, 53].

The dynamics for the target agent are given by

$$\ddot{q}_0 = f_0(q_0, \dot{q}_0, t), \tag{4-2}$$

where $q_0, \dot{q}_0, \ddot{q}_0 \in \mathbb{R}^n$ denote the target's unknown generalized position, velocity, and acceleration, respectively, and the function $f_0 : \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^n$ is unknown, of class C^3, and the mappings $t \mapsto \nabla f_0(q_0, \dot{q}_0, t)$ and $t \mapsto \nabla^2 f_0(q_0, \dot{q}_0, t)$ are uniformly bounded.

Assumption 4.1. *There exist known constants* $\overline{q}_0, \overline{\dot{q}}_0, \overline{\ddot{q}}_0, \overline{\dddot{q}}_0 \in \mathbb{R}_{>0}$ *such that* $\|q_0(t)\| \leq \overline{q}_0$, $\|\dot{q}_0(t)\| \leq \overline{\dot{q}}_0$, $\|\ddot{q}_0(t)\| \leq \overline{\ddot{q}}_0$, *and* $\|\dddot{q}_0(t)\| \leq \overline{\dddot{q}}_0$ *for all* $t \in [0, \infty)$.

4.1.2 Control Objective

Each agent $i \in \mathcal{V}$ can measure the relative position $q_{i,j} \in \mathbb{R}^n$ and relative velocity $\dot{q}_{i,j} \in \mathbb{R}^n$ between itself and any agent $j \in \overline{\mathcal{N}_i}$, where the relative position is defined as

$$q_{ij} \triangleq q_j - q_i. \tag{4-3}$$

The objective is to design a distributed controller that enables all agents to track the target using only relative state measurements, given the unknown dynamics described by (4-1) and (4-2). To quantify the tracking performance, the target tracking error $e_i \in \mathbb{R}^n$ is defined as

$$e_i \triangleq q_0 - q_i. \tag{4-4}$$

Using (4-4), the neighborhood tracking error $\eta_i \in \mathbb{R}^n$ is defined as

$$\eta_i \triangleq b_i e_i + \sum_{j \in \mathcal{N}_i} a_{ij} q_{ij}, \tag{4-5}$$

where $b_i \in \{0, 1\}$ indicates whether agent $i \in \mathcal{V}$ senses the target. Using (4-3), (4-5) is expressed in an equivalent analytical form as

$$\eta_i = b_i e_i - \sum_{j \in \mathcal{N}_i} a_{ij} \left(e_j - e_i \right). \tag{4-6}$$

4.2 Control Design

Let $\mathcal{I} \triangleq [t_0, t_1]$ denote the interval where solutions exist for the subsequent closed-loop error system, where $t_0, t_1 \in \mathbb{R}_{\geq 0}$ and $t_1 > t_0$. Based on the subsequent stability analysis, the continuous control input u_i for each agent $i \in \mathcal{V}$ is designed as

$$u_i = g_i^+ \left(q_i, \dot{q}_i, t \right) \left(k_3 \dot{\eta}_i + \left(\left(k_1 + k_2 \right) k_3 + 1 \right) \eta_i + \left(k_1 + k_2 \right) b_i \dot{e}_i + \left(1 + k_1 k_2 \right) b_i e_i + \hat{v}_i \right) \tag{4-7}$$

where $\hat{v}_i \in \mathbb{R}^n$ is designed as a Filippov solution to

$$\dot{\hat{v}}_i = \left(k_1 + \left(1 + k_1 k_2 \right) k_3 \right) \eta_i + k_4 \text{sgn} \left(\dot{\eta}_i + k_1 \eta_i \right), \tag{4-8}$$

where $k_1, k_2, k_3, k_4 \in \mathbb{R}_{>0}$ are user-defined constants. To aid in the stability analysis, the graph interaction matrix $\mathcal{H} \in \mathbb{R}^{nN \times nN}$ is defined as

$$\mathcal{H} \triangleq (L + B) \otimes I_n, \tag{4-9}$$

where $B \triangleq \text{diag}\{b_1, \ldots, b_N\} \in \mathbb{R}^{N \times N}$. Using (4-9), the neighborhood position error in (4-6) is expressed in an ensemble representation as

$$\eta = \mathcal{H} e, \tag{4-10}$$

where $\eta \triangleq (\eta_i)_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$ and $e \triangleq (e_i)_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Based on the subsequent stability analysis, define the filtered tracking error $r_1 \in \mathbb{R}^{nN}$ as

$$r_1 \triangleq \dot{e} + k_1 e. \tag{4-11}$$

Using (4-10), (4-11) is expressed as

$$\mathcal{H} r_1 = \dot{\eta} + k_1 \eta. \tag{4-12}$$

Similarly, define the auxiliary tracking error $r_2 \in \mathbb{R}^{nN}$ as

$$r_2 \triangleq \dot{r}_1 + k_2 r_1 + e. \tag{4-13}$$

Using (4-12), (4-13) is expressed as

$$\mathcal{H} r_2 = \ddot{\eta} + (k_1 + k_2)\, \dot{\eta} + (1 + k_1 k_2)\, \eta. \tag{4-14}$$

Using (4-8), (4-12), and (4-14), the time-derivative of (4-7) is expressed in an ensemble representation as

$$\dot{u} = g^{+}\left(\mathcal{H}\left(r_1 + k_3 r_2\right) + k_4 \text{sgn}\left(\mathcal{H} r_1\right) + \mathcal{B}\left(k_1 \ddot{e} + k_2 \dot{r}_1 + \dot{e}\right)\right), \tag{4-15}$$

where $u \triangleq [u_i]_{i\in\mathcal{V}} \in \mathbb{R}^{N\sum_{i\in\mathcal{V}} m_i}$, $g^{+} \triangleq \text{diag}\left\{g_1^{+}, \ldots, g_N^{+}\right\} \in \mathbb{R}^{N\left(\sum_{i\in\mathcal{V}} m_i \times n\right)}$, and $\mathcal{B} \triangleq B \otimes I_n \in \mathbb{R}^{nN}$. Substituting (4-15) into the time-derivative of (4-13) yields

$$\begin{aligned}\dot{r}_2 = {} & h_B + \widetilde{h} - k_4 \text{sgn}\left(\mathcal{H} r_1\right) - \left(k_3 \mathcal{H} + \left(k_1 + k_2\right)\left(\mathcal{B} - I_{nN}\right)\right) r_2 \\ & - \left(\mathcal{H} + \left(1 - 2k_1^2 - k_1 k_2\right)\left(\mathcal{B} - I_{nN}\right)\right) r_1 + \left(k_1\left(2 - k_1^2\right) + k_2\right)\left(\mathcal{B} - I_{nN}\right) e,\end{aligned} \tag{4-16}$$

where

$$\begin{aligned}h_B \triangleq {} & \left[\dot{f}_0\left(q_0, \dot{q}_0, t\right) - \dot{f}_i\left(q_0, \dot{q}_0, t\right) - \dot{d}_i\left(t\right)\right]_{i\in\mathcal{V}} \\ & - \left[\dot{g}_i\left(q_0, \dot{q}_0, t\right) g_i^{+}\left(q_0, \dot{q}_0, t\right)\left(\ddot{q}_0 - f_i\left(q_0, \dot{q}_0, t\right) - d_i\left(t\right)\right)\right]_{i\in\mathcal{V}},\end{aligned}$$

and

$$\begin{aligned}\widetilde{h} \triangleq {} & \left[\dot{f}_i\left(q_0, \dot{q}_0, t\right) - \dot{f}_i\left(q_i, \dot{q}_i, t\right) + \dot{g}_i\left(q_0, \dot{q}_0, t\right) g_i^{+}\left(q_0, \dot{q}_0, t\right)\left(\ddot{q}_0 - f_i\left(q_0, \dot{q}_0, t\right) - d_i\left(t\right)\right)\right]_{i\in\mathcal{V}} \\ & - \left[\dot{g}_i\left(q_i, \dot{q}_i, t\right) g_i^{+}\left(q_i, \dot{q}_i, t\right)\left(\ddot{q}_i - f_i\left(q_i, \dot{q}_i, t\right) - d_i\left(t\right)\right)\right]_{i\in\mathcal{V}}.\end{aligned}$$

Assumption 4.2. *The graph G is connected with at least one $b_i > 0$, for all $i \in \mathcal{V}$.*

Lemma 4.1. *[81, Remark 1.2] Following Assumption 4.2, $\mathcal{H}$ is positive definite.*

Define the concatenated state vector $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^{3nN}$ as $z \triangleq \begin{bmatrix} e^\top & r_1^\top & r_2^\top \end{bmatrix}^\top$. Using (4-11), (4-13), and (4-16), the closed-loop error system is expressed as

$$\dot{z} = \begin{bmatrix} r_1 - k_1 e \\ r_2 - k_2 r_1 - e \\ \begin{pmatrix} h_B + \widetilde{h} - k_4 \text{sgn}\left(\mathcal{H} r_1\right) - \left(k_3 \mathcal{H} + \left(k_1 + k_2\right)\left(\mathcal{B} - I_{nN}\right)\right) r_2 \\ -\left(\mathcal{H} + \left(1 - 2k_1^2 - k_1 k_2\right)\left(\mathcal{B} - I_{nN}\right)\right) r_1 + \left(k_1\left(2 - k_1^2\right) + k_2\right)\left(\mathcal{B} - I_{nN}\right) e, \end{pmatrix} \end{bmatrix}. \tag{4-17}$$

The following section provides a Lyapunov-based stability analysis to provide exponential tracking error convergence guarantees with the developed controller. Some supporting lemmas are first presented to facilitate the subsequent analysis.

Lemma 4.2. *There exist known constants $\chi_1, \chi_2 \in \mathbb{R}_{>0}$ such that $\|h_B\| \leq \chi_1$ and $\left\|\dot{h}_B\right\| \leq \chi_2$.*

Proof. By Assumption 4.1, q_0 is defined over a compact set $\mathcal{Q} \subset \mathbb{R}^n$, where $\mathcal{Q} \triangleq \left\{\iota \in \mathbb{R}^n : \|\iota\| \leq \overline{q}_0\right\}$. Using the definition of h_B and the triangle inequality yields $\|h_B\| \leq N \left\| \frac{\mathrm{d}}{\mathrm{d}t} f_0\left(q_0, \dot{q}_0, t\right) - \frac{\mathrm{d}}{\mathrm{d}t} f_i\left(q_0, \dot{q}_0, t\right) - \frac{\mathrm{d}}{\mathrm{d}t} d_i\left(t\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(g_i\left(q_0, \dot{q}_0, t\right)\right) g_i^+\left(q_0, \dot{q}_0, t\right)\left(f_i\left(q_0, \dot{q}_0, t\right) + d_i\left(t\right) - \ddot{q}_0\right)\right\|$. Applying Assumption 4.1, the chain-rule, the Cauchy-Schwarz inequality, and the triangle inequality yields

$$\begin{aligned} \|h_B\| \leq N \left(\overline{\dot{q}}_0 \left(\left\| \frac{\partial f_0}{\partial q_0} \right\| + \left\| \frac{\partial f_i}{\partial q_0} \right\| \right) + \overline{\ddot{q}}_0 \left(\left\| \frac{\partial f_0}{\partial \dot{q}_0} \right\| + \left\| \frac{\partial f_i}{\partial \dot{q}_0} \right\| \right) + \left\| \frac{\partial f_0}{\partial t} \right\| + \left\| \frac{\partial f_i}{\partial t} \right\| + \overline{\dot{d}}_i \right) \\ + N \left(\left\| \frac{\partial g_i}{\partial q_0} \right\| \overline{\dot{q}}_0 + \left\| \frac{\partial g_i}{\partial \dot{q}_0} \right\| \overline{\ddot{q}}_0 + \left\| \frac{\partial g_i}{\partial t} \right\| \right) \left\| g_i^+\left(q_0, \dot{q}_0, t\right) \right\| \left(\left\| f_i\left(q_0, \dot{q}_0, t\right) \right\| + \overline{d}_i + \overline{\ddot{q}}_0 \right). \end{aligned}$$

Since q_0 is defined over the compact set $\mathcal{Q}$, applying the mean value theorem gives $\|h_B\| \leq \chi_1$.

Using the definition of h_B and the triangle inequality yields $\left\|\dot{h}_B\right\| \leq N \left\| \frac{\mathrm{d}^2}{\mathrm{d}t^2} f_0\left(q_0, \dot{q}_0, t\right) - \frac{\mathrm{d}^2}{\mathrm{d}t^2} f_i\left(q_0, \dot{q}_0, t\right) - \frac{\mathrm{d}^2}{\mathrm{d}t^2} d_i\left(t\right) - \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(g_i\left(q_0, \dot{q}_0, t\right)\right) g_i^+\left(q_0, \dot{q}_0, t\right)\left(f_i\left(q_0, \dot{q}_0, t\right) + d_i\left(t\right) - \ddot{q}_0\right)\right\|$. Applying Assumption 4.1, the chain-rule, the product rule, the Cauchy-Schwarz inequality, and

the triangle inequality yields

$$\begin{aligned}\left\|\dot{h}_B\right\| \leq N & \left(\overline{\dot{q}}_0\left(\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_0}{\partial q_0}\right)\right\|+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_i}{\partial q_0}\right)\right\|\right)+\overline{\ddot{q}}_0\left(\left\|\frac{\partial f_0}{\partial q_0}\right\|+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_0}{\partial \dot{q}_0}\right)\right\|+\left\|\frac{\partial f_i}{\partial q_0}\right\|+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_i}{\partial \dot{q}_0}\right)\right\|\right)\right) \\ & + N\left(\overline{\ddot{q}}_0\left(\left\|\frac{\partial f_0}{\partial \dot{q}_0}\right\|+\left\|\frac{\partial f_i}{\partial \dot{q}_0}\right\|\right)+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_0}{\partial t}\right)\right\|+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_i}{\partial t}\right)\right\|+\overline{\dddot{d}}_i\right) \\ & + N\left(\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial g_i}{\partial q_0}\right)\right\|\overline{\dot{q}}_0+\overline{\ddot{q}}_0\left(\left\|\frac{\partial g_i}{\partial q_0}\right\|+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial g_i}{\partial \dot{q}_0}\right)\right\|\right)+\left\|\frac{\partial g_i}{\partial \dot{q}_0}\right\|\overline{\ddot{q}}_0+\left\|\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial g_i}{\partial t}\right)\right\|\right)\left\|g_i^{+}\right\|\left(\|f_i\|+\overline{d}_i+\overline{\ddot{q}}_0\right).\end{aligned}$$

Since q_0 is defined over the compact set $\mathcal{Q}$, applying the mean value theorem gives $\left\|\dot{h}_B\right\| \leq \chi_2$. □

Remark 4.2. If uncertainty bounds χ_1, χ_2 are unknown, adaptive RISE methods (e.g., [3]) can estimate related bounds online. This typically involves a trade-off, as such adaptive approaches often yield asymptotic stability or ultimate boundedness, contrasting with the exponential stability targeted herein, which generally remains an open challenge for adaptive RISE.

Lemma 4.3. *[29, Lemma 5] There exists a strictly increasing function* $\rho : [0, \infty) \to [0, \infty)$ *such that* $\left\|\widetilde{h}\right\| \leq \rho\left(\|z\|\right)\|z\|$.

4.3 Stability Analysis

To establish exponential stability for the closed-loop error system in (4-17) a P-function is introduced. This function is used to develop a strict Lyapunov function and is designed to be non-negative under specific gain conditions. The P-function $P : \mathcal{I} \to \mathbb{R}_{\geq 0}$ is designed as

$$P \triangleq k_4\left\|\mathcal{H} r_1\right\|_1 - r_1^{\top}\mathcal{H}^{\top} h_B + \mathrm{e}^{-\lambda_P t} * \left(r_1^{\top}\mathcal{H}^{\top}\dot{h}_B + \left(k_2 - \lambda_P\right)\left(k_4\left\|\mathcal{H} r_1\right\|_1 - r_1^{\top}\mathcal{H}^{\top} h_B\right)\right), \quad (4\text{-}18)$$

where $\lambda_P \in \mathbb{R}_{>0}$ is a user-defined constant, and the symbol '$*$' denotes the convolutional integral, defined for any given function $\alpha(t) : \mathcal{I} \to \mathbb{R}$ as $\mathrm{e}^{-\lambda_P t} * \alpha(t) = \int_{t_0}^{t} \mathrm{e}^{-\lambda_P(t-\tau)}\alpha(t)\,\mathrm{d}\tau$. Using Leibniz's rule, the time derivative of a convolution integral satisfies
$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-\lambda_P t} * \alpha\right) = \alpha(t) - \lambda_P \int_{t_0}^{t} \mathrm{e}^{-\lambda_P(t-\tau)}\alpha(\tau)\,\mathrm{d}t$ which simplifies to
$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-\lambda_P t} * \alpha\right) = \alpha(t) - \lambda_P \mathrm{e}^{-\lambda_P t} * \alpha(t)$. Using (4-18) and the definition of the convolutional integral yields

$$P(t_0) = k_4\left\|\mathcal{H} r_1(t_0)\right\|_1 - r_1^{\top}(t_0)\,\mathcal{H}^{\top} h_B(t_0). \quad (4\text{-}19)$$

Since $t \mapsto \mathcal{H}r_1(t)$ is absolutely continuous and $\|\cdot\|_1$ is globally Lipschitz, the mapping $t \mapsto \|r_1(t)\|_1$ is differentiable almost everywhere. By the chain rule from [90, Theorem 2.2], the derivative of $\|\mathcal{H}r_1\|_1$ is given almost everywhere by

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathcal{H}r_1\|_1 = \frac{\mathrm{d}}{\mathrm{d}t}\|\dot{\eta} + k_1\eta\|_1 \overset{\text{a.e.}}{\in} (\ddot{\eta} + k_1\dot{\eta})^\top K\,[\mathrm{sgn}]\,(\dot{\eta} + k_1\eta)\,.$$

By the definition of the L_1-norm,

$$\|\dot{\eta} + k_1\eta\|_1 = (\dot{\eta} + k_1\eta)^\top K\,[\mathrm{sgn}]\,(\dot{\eta} + k_1\eta)\,.$$

Using the definitions of the chain rule and the L_1-norm, taking the time-derivative of (4-18), using Leibniz's rule, and substituting (4-12), (4-14), and (4-18) into the resulting expression yields that $t \mapsto P(t)$ satisfies the differential inclusion

$$\dot{P} \overset{\text{a.e.}}{\in} -\lambda_P P - r_2^\top \mathcal{H}^\top (h_B - k_4 K\,[\mathrm{sgn}]\,(\mathcal{H}r_1))\,. \tag{4-20}$$

To facilitate the inclusion of the P-function into the subsequent Lyapunov function candidate, P must be designed to be non-negative under certain sufficient gain conditions.

Proposition 4.1. *For the P-function defined as in (4-18), if $k_4 > \chi_1 + \frac{\chi_2}{k_2 - \lambda_P}$ with $\lambda_P \in (0, k_2)$, then $P(t) \geq 0$, for all $t \in \mathcal{I}$.*

Proof. By Lemma 4.2, the $k_4\|\mathcal{H}r_1\|_1 - r_1^\top \mathcal{H}^\top h_B$ term in (4-18) satisfies

$$k_4\|\mathcal{H}r_1\|_1 - r_1^\top \mathcal{H}^\top h_B \geq (k_4 - \chi_1)\|\mathcal{H}r_1\|_1 \geq 0,$$

by the Cauchy-Schwarz inequality and the condition $k_4 > \chi_1$. The convolution integrand in (4-18) is bounded as

$$r_1^\top \mathcal{H}^\top \dot{h}_B + (k_2 - \lambda_P)\left(k_4 \|\mathcal{H} r_1\|_1 - r_1^\top \mathcal{H}^\top h_B\right) \geq \|\mathcal{H} r_1\|_1 \left((k_2 - \lambda_P)(k_4 - \chi_1) - \chi_2\right).$$

This expression is non-negative when $(k_2 - \lambda_P)(k_4 - \chi_1) \geq \chi_2$ which is equivalent to $k_4 \geq \chi_1 + \frac{\chi_2}{k_2 - \lambda_P}$. Since $k_4 > \chi_1 + \frac{\chi_2}{k_2 - \lambda_P}$ by assumption, the convolution integrand is positive, making the entire convolution positive. Therefore, $P(t) \geq 0$, for all $t \in \mathcal{I}$. □

To state the main results, the following definitions are introduced. Let $W : \mathbb{R}^{3nN} \to \mathbb{R}_{\geq 0}$ be defined as

$$W(\sigma) \triangleq \underline{\lambda}_Q^{-\frac{1}{2}} \sqrt{\overline{\lambda}_Q \|\sigma\|^2 + 2(k_4 + \chi_1)\|\mathcal{H}\|_1 \|\sigma\|_1}, \tag{4-21}$$

for all $\sigma \in \mathbb{R}^{3nN}$, where $Q \triangleq \operatorname{diag}(I_{2nN}, \mathcal{H}) \in \mathbb{R}^{3nN \times 3nN}$ and χ_1 is defined as in Lemma 4.2. Additionally, let $k_{\min} \in \mathbb{R}$ be a constant gain defined as
$k_{\min} \triangleq \min\{k_1 - \frac{1}{2}, k_2 - \frac{1}{2}, 2(k_1 + k_2)\left(\underline{\lambda}_{\mathcal{H}} \underline{\lambda}_{(\mathcal{B} - I_{nN})} + k_3 \underline{\lambda}_{\mathcal{H}}^2\right) -$
$\left(\overline{\lambda}_{(I_{nN} - \mathcal{H}^2)} + \left(1 + 2k_1^2 + k_1 k_2\right)\overline{\lambda}_{\mathcal{H}} \overline{\lambda}_{(\mathcal{B} - I_{nN})}\right)^2 - \left(k_1\left(2 - k_1^2\right) + k_2\right)^2 \overline{\lambda}_{\mathcal{H}}^2 \overline{\lambda}_{(\mathcal{B} - I_{nN})}^2\}$, $\lambda_V \in \mathbb{R}_{>0}$ be the desired rate of convergence, and the set of stabilizing initial conditions $\mathcal{S} \subset \mathbb{R}^{3nN}$ be defined as

$$\mathcal{S} \triangleq \left\{\sigma \in \mathbb{R}^{3nN} : \rho(W(\sigma)) \leq \frac{k_{\min} - \lambda_V}{\overline{\lambda}_{\mathcal{H}}}\right\}. \tag{4-22}$$

Let $\xi : \mathcal{I} \to \mathbb{R}^{3nN+1}$ be defined as $\xi(t) = \begin{bmatrix} z^\top(t) & P(t) \end{bmatrix}^\top$, and $\psi : \mathbb{R}^{3nN+1} \times \mathbb{R}_{\geq 0} \rightrightarrows \mathbb{R}^{3nN+1}$ denote the set-valued map

$$\psi(\xi, t) \triangleq \begin{bmatrix} r_1 - k_1 e \\ r_2 - k_2 r_1 - e \\ \left(\begin{array}{c} h_B + \widetilde{h} - k_4 K [\text{sgn}] (\mathcal{H} r_1) - (k_1 + k_2)(\mathcal{B} - I_{nN}) r_2 \\ -\left(1 - 2k_1^2 - k_1 k_2\right)(\mathcal{B} - I_{nN}) r_1 - \mathcal{H} r_1 - k_3 \mathcal{H} r_2 \\ +\left(k_1\left(2 - k_1^2\right) + k_2\right)(\mathcal{B} - I_{nN}) e \end{array} \right) \\ -\lambda_P P - r_2^\top \mathcal{H}^\top h_B + k_4 r_2^\top \mathcal{H}^\top K [\text{sgn}] (\mathcal{H} r_1) \end{bmatrix}. \tag{4-23}$$

Then using (4-17) and (4-20), it follows that the trajectories $t \mapsto \xi(t)$ satisfy the differential inclusion $\dot{\xi} \overset{\text{a.e.}}{\in} \psi(\xi, t)$.

Theorem 4.1. *All solutions to (4-17) with $z(t_0) \in int(\mathcal{S})$ satisfy $\|z(t)\| \leq W(z(t_0))\, e^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)}$, for all $t \in [t_0, \infty)$, provided that the sufficient control gains k_1, k_2, k_3 are selected to satisfy $k_1 > \frac{1}{2}$, $k_2 > \frac{1}{2}$, $k_3 > \left(\overline{\lambda}_{(I_{nN} - \mathcal{H}^2)} + \overline{\lambda}_{\mathcal{H}} \overline{\lambda}_{(\mathcal{B} - I_{nN})}\right)^2 + 2\overline{\lambda}_{\mathcal{H}}^2 \overline{\lambda}_{(\mathcal{B} - I_{nN})}^2$, k_4 is selected to satisfy Proposition 4.1, and Assumptions 4.1 and 4.2 hold.*

Proof. Consider the function $V : \mathbb{R}^{3nN+1} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined as

$$V(\xi) \triangleq \frac{1}{2} z^\top Q z + P. \tag{4-24}$$

For the subsequent analysis, consider all arbitrary trajectories $t \mapsto \xi(t)$ satisfying (4-19) and $z(t_0) \in \text{int}(\mathcal{S})$. Since Proposition 4.1 holds, it follows that these trajectories satisfy $P(t) \geq 0$ for all $t \in \mathcal{I}$ which implies that $V(\xi(t)) > 0$ for all $t \in \mathcal{I}$. Invoking the Rayleigh quotient theorem and using (4-24) yields

$$\frac{\lambda_Q}{2} \|z\|^2 + P \leq V(\xi). \tag{4-25}$$

Based on the chain rule for differential inclusions in [90, Theorem 2.2], the derivative of $t \mapsto V(\xi(t))$ exists almost everywhere and is a solution to $\dot{V}(\xi) \overset{\text{a.e.}}{\in} \dot{\tilde{V}}(\xi)$, where the set $\dot{\tilde{V}}(\xi)$ is

defined as $\dot{\widetilde{V}}(\xi) \triangleq \bigcap_{\zeta \in \partial V(\xi)} \zeta^\top \psi(\xi, t)$. Since $V(\xi)$ is continuously differentiable for all $\xi \in \mathbb{R}^{3nN+1}$, Clarke's gradient reduces to the singleton $\{\nabla V(\xi)\} = \left\{ \begin{bmatrix} e^\top & r_1^\top & r_2^\top \mathcal{H} & 1 \end{bmatrix}^\top \right\}$. Thus, $\dot{\widetilde{V}}(\xi) = \bigcap_{\zeta \in \partial V(\xi)} \zeta^\top \psi(\xi, t) = (\nabla V(\xi))^\top \psi(\xi, t)$. Evaluating $\dot{\widetilde{V}}(\xi)$ yields

$$\begin{aligned}\dot{\widetilde{V}}(\xi) = &-k_1 e^\top e - k_2 r_1^\top r_1 - \lambda_P P + r_2^\top \mathcal{H} \widetilde{h} - r_2^\top \left((k_1 + k_2) \mathcal{H} (\mathcal{B} - I_{nN}) + k_3 \mathcal{H}^2 \right) r_2 \\ &+ r_2^\top \left(I_{nN} - \mathcal{H}^2 - \left(1 - 2k_1^2 - k_1 k_2 \right) \mathcal{H} (\mathcal{B} - I_{nN}) \right) r_1 + \left(k_1 \left(2 - k_1^2 \right) + k_2 \right) r_2^\top \mathcal{H} (\mathcal{B} - I_{nN}) e \\ &+ k_4 r_2^\top \mathcal{H}^\top \left(K[\text{sgn}](\mathcal{H} r_1) - K[\text{sgn}](\mathcal{H} r_1) \right). \end{aligned} \tag{4-26}$$

By [33, Footnote 2], the right-hand side of (4-26) is continuous almost everywhere. Specifically, it is discontinuous only on a set of times with Lebesgue measure zero, where $K[\text{sgn}](\mathcal{H} r_1) - K[\text{sgn}](\mathcal{H} r_1) \neq 0$. Since this set is Lebesgue negligible, $K[\text{sgn}](\mathcal{H} r_1) \overset{\text{a.e.}}{=} \{\text{sgn}(\mathcal{H} r_1)\}$. Furthermore, from Lemma 4.3, $\left\| \widetilde{h} \right\| \leq \rho(\|z\|) \|z\|$. Applying the Cauchy-Schwarz inequality, the triangle inequality, and Young's inequality yields that (4-26) is upper bounded as

$$\begin{aligned}\dot{\widetilde{V}}(\xi) \overset{\text{a.e.}}{\leq} &-\left(k_1 - \frac{1}{2} \right) \|e\|^2 - \left(k_2 - \frac{1}{2} \right) \|r_1\|^2 - (k_1 + k_2) \left(\underline{\lambda}_{\mathcal{H}} \underline{\lambda}_{(\mathcal{B} - I_{nN})} + k_3 \underline{\lambda}_{\mathcal{H}}^2 \right) \|r_2\|^2 \\ &+ \left(\overline{\lambda}_{(I_{nN} - \mathcal{H}^2)} + \left(1 + 2k_1^2 + k_1 k_2 \right) \overline{\lambda}_{\mathcal{H}} \overline{\lambda}_{(\mathcal{B} - I_{nN})} \right)^2 \|r_2\|^2 \\ &+ \frac{1}{2} \left(k_1 \left(2 - k_1^2 \right) + k_2 \right)^2 \overline{\lambda}_{\mathcal{H}}^2 \overline{\lambda}_{(\mathcal{B} - I_{nN})}^2 \|r_2\|^2 \\ &+ \overline{\lambda}_{\mathcal{H}} \rho(\|z\|) \|z\|^2 - \lambda_P P. \end{aligned} \tag{4-27}$$

Since k_1, k_2, and k_3 are selected to satisfy $k_1 > \frac{1}{2}$, $k_2 > \frac{1}{2}$, and $k_3 > \left(\overline{\lambda}_{(I_{nN} - \mathcal{H}^2)} + \overline{\lambda}_{\mathcal{H}} \overline{\lambda}_{(\mathcal{B} - I_{nN})} \right)^2 + 2 \overline{\lambda}_{\mathcal{H}}^2 \overline{\lambda}_{(\mathcal{B} - I_{nN})}^2$, it follows that $k_{\min} > 0$. Consequently, using the definition of $\|z\|$ yields that (4-27) is upper-bounded as

$$\dot{\widetilde{V}}(\xi) \overset{\text{a.e.}}{\leq} -\left(k_{\min} - \overline{\lambda}_{\mathcal{H}} \rho(\|z\|) \right) \|z\|^2 - \lambda_P P. \tag{4-28}$$

Since $z(t_0) \in \mathcal{S}$, it follows from (4-22) that $k_{\min} > \lambda_V + \overline{\lambda}_{\mathcal{H}} \rho(W(z(t_0)))$. Therefore,

$$\dot{\tilde{V}}(\xi) \overset{\text{a.e.}}{\leq} -\left(\lambda_V + \overline{\lambda}_{\mathcal{H}}(\rho(W(z(t_0))) - \rho(\|z\|))\right)\|z\|^2 - \lambda_P P. \tag{4-29}$$

By the definition of W in (4-21), it follows that $\|z(t_0)\| \leq W(z(t_0))$. Because the solution $t \mapsto z(t)$ is continuous, z cannot instantaneously escape $\mathcal{S}$ at t_0. Therefore, there exists a time interval $\mathcal{I}_{\mathcal{S}}$ satisfying $\mathcal{I}_{\mathcal{S}} \subseteq \mathcal{I}$ such that $z(t) \in \mathcal{S}$ for all $t \in \mathcal{I}_{\mathcal{S}}$, implying $\|z(t)\| < W(z(t_0))$. Because ρ is strictly increasing, $\rho(\|z(t)\|) < \rho(W(z(t_0)))$ for all $t \in \mathcal{I}_{\mathcal{S}}$. Consequently, using (4-25), selecting $\lambda_P > \lambda_V$ and recalling $\dot{V}(\xi) \overset{\text{a.e.}}{\in} \dot{\tilde{V}}(\xi)$ yields that

$$\dot{V}(\xi) \leq -\frac{2\lambda_V}{\overline{\lambda}_Q} V(\xi), \tag{4-30}$$

for all $t \in \mathcal{I}_{\mathcal{S}}$. Solving the differential inequality in (4-30) and using (4-25) yields that $\|z\|$ is upper-bounded as

$$\|z(t)\| \leq \sqrt{2\underline{\lambda}_Q^{-1} V(\xi(t_0))}\, \mathrm{e}^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)}, \tag{4-31}$$

for all $t \in \mathcal{I}_{\mathcal{S}}$. Using (4-19) yields that
$V(\xi(t_0)) = \frac{1}{2}\overline{\lambda}_Q \|z(t_0)\|^2 + k_4 \|\mathcal{H} r_1(t_0)\|_1 - r_1^\top(t_0)\mathcal{H}^\top h_B(t_0)$. Since h_B is bounded by Lemma 4.2, using the Cauchy-Schwarz inequality and the fact that $\|r_1(t_0)\| \leq \|r_1(t_0)\|_1 \leq \|z(t_0)\|_1$ yields

$$\sqrt{2\underline{\lambda}_Q^{-1} V(\xi(t_0))} \leq W(z(t_0)), \tag{4-32}$$

where W is defined in (4-21). Finally, applying (4-32) to (4-31) yields
$\|z(t)\| \leq W(z(t_0))\, \mathrm{e}^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)}$, for all $t \in \mathcal{I}_{\mathcal{S}}$.

It remains to be shown that $\mathcal{I}_{\mathcal{S}}$ can be extended to $[t_0, \infty)$. Let $t \mapsto \xi(t)$ be a maximal solution to the differential inclusion $\dot{\xi} \overset{\text{a.e.}}{\in} \psi(\xi, t)$ with initial conditions satisfying (4-19) and $z(t_0) \in \text{int}(\mathcal{S})$. From the preceding analysis, $z(t) \in \text{int}(\mathcal{S})$ for all $t \in \mathcal{I}_{\mathcal{S}}$. This implies that $\xi(t) \in \mathcal{D} \triangleq \left\{\mathfrak{S} \in \mathbb{R}^{3nN+1} : \mathfrak{S} = \begin{bmatrix} \sigma^\top & \varsigma \end{bmatrix}^\top, \sigma \in \text{int}(\mathcal{S}), \varsigma \in \mathbb{R}_{\geq 0}\right\}$. For any compact

subinterval $\mathcal{J} \subseteq \mathcal{I}_{\mathcal{S}}$, $\|z(t)\| \leq W(z(t_0))\, e^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)}$, for all $t \in \mathcal{J}$. Since W is a locally bounded function and $e^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)} \leq 1$ for all $t \geq t_0$, the mapping $t \mapsto z(t)$ is uniformly bounded on $\mathcal{J}$. By definition, $(\xi, t) \mapsto \psi(\xi, t)$ is locally bounded when ξ is bounded. Since $z(t)$ is bounded on $\mathcal{J}$ and $P(t)$ is non-negative, the trajectory $t \mapsto (\xi(t), t)$ is precompact. Furthermore, since the map $(\xi, t) \mapsto \psi(\xi, t)$ is locally bounded, it follows from [27, Remark 3.4] that $\bigcup_{t\in\mathcal{J}} \psi(\xi(t), t)$ is bounded for every compact subinterval $\mathcal{J} \subseteq \mathcal{I}_{\mathcal{S}}$. Therefore, the conditions of [27, Lemma 3.3] are satisfied, guaranteeing that $t \mapsto \xi(t)$ is complete, i.e., $\mathcal{I} = [t_0, \infty)$. Thus, all trajectories satisfying $z(t_0) \in \text{int}(\mathcal{S})$ also satisfy $\|z(t)\| \leq W(z(t_0))\, e^{-\frac{2\lambda_V}{\overline{\lambda}_Q}(t-t_0)}$, for all $t \in [t_0, \infty)$.

Recall that $k_{\min} > \lambda_V + \overline{\lambda}_{\mathcal{H}}\rho(W(z(t_0)))$, implying that the exponential stability result is semi-global [76, Remark 2], as the set of stabilizing initial conditions in (4-22) can be made arbitrarily large by appropriately adjusting $k_{\min}$ to encompass any $z(t_0) \in \mathbb{R}^{3nN}$.

Because k_4, χ_1, and λ_V are independent of the initial time t_0 or initial condition $z(t_0)$, the exponential convergence is uniform [51]. Additionally, the convergence and boundedness of $\|z\|$ implies the convergence and boundedness of $\|e\|$, $\|r_1\|$, and $\|r_2\|$. Therefore, since $q_0, \dot{q}_0, \ddot{q}_0 \in \mathcal{L}_\infty$ by Assumption 4.1, using (4-4), (4-6) and (4-10)-(4-14) yields that $q_i, \dot{q}_i, \ddot{q}_i \in \mathcal{L}_\infty$, for all $i \in \mathcal{V}$. Thus, $f_i(q_i, \dot{q}_i, t)$ and $g_i(q_i, \dot{q}_i, t)$ are bounded for all $i \in \mathcal{V}$. Therefore, using (4-1) yields that $u \in \mathcal{L}_\infty$. □

4.4 Simulation

Numerical simulations were performed to evaluate the proposed distributed controller. The scenario involved $N = 8$ agents tracking a dynamic target agent, with inter-agent communication following a cycle graph topology. The pinning matrix was $B = \text{diag}\{1, 0, 1, 0, 1, 0, 1, 0\}$. Agent initial positions were sampled from $U(-10, 10)$ m per axis, with zero initial velocities. The target's initial position was sampled from $U(-10, 10)$ m with an initial velocity of

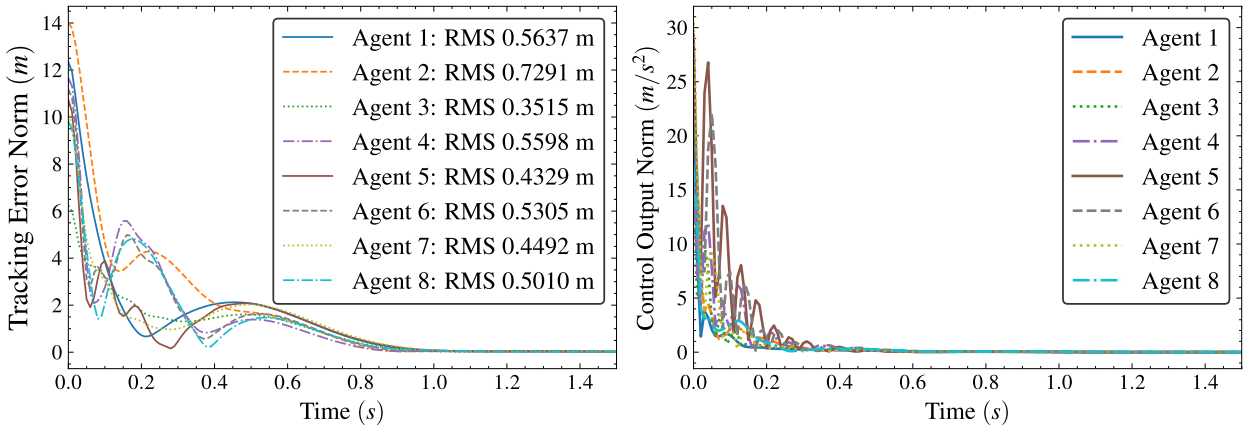

Figure 4-1. Left: Tracking error norm for $t \in [0, 1.5]$. Right: Control effort norm for $t \in [0, 1.5]$s.

$\begin{bmatrix} 1 & -1 & 0.5 \end{bmatrix}^\top m/s$. Each agent i adhered to the dynamics in (4-1) with

$$f_i = \begin{bmatrix} c_{i,1}\,(y_i - z_i) + c_{i,2}\tanh(\dot{x}_i t) \\ c_{i,3}\,(z_i - x_i) + c_{i,4}\tanh(\dot{y}_i t) \\ c_{i,5}\,(x_i - y_i) + c_{i,6}\tanh(\dot{z}_i t) \end{bmatrix},$$

$$g_i = \begin{bmatrix} 1 - c_{i,7}\cos(t) & 0 & 0 \\ 0 & 1 - c_{i,8}\sin(t) & 0 \\ 0 & 0 & 1 - c_{i,9}\cos(t)\sin(t) \end{bmatrix},$$

and

$$d_i = \begin{bmatrix} c_{i,10}\cos(t) \\ c_{i,11}\sin(t) \\ c_{i,12}\cos(t)\sin(t) \end{bmatrix}$$

with parameters $c_{i,j}$ sampled from $U(-0.5, 0.5)$. The target's dynamics were

$$f_i = \begin{bmatrix} \sin(x_0) - \cos(y_0\dot{x}_0) \\ \cos(z_0\dot{y}_0) - \sin(x_0) \\ -\sin(y_0\dot{z}_0) - \sin(z_0) \end{bmatrix}.$$

Relative position and velocity measurements were corrupted by additive zero-mean Gaussian noise ($\sigma = 0.001$ m and $\sigma = 0.001$ m/s, respectively). Furthermore, heterogeneous time delays, unique to each agent, were introduced: communication delays for inter-agent information, actuation delays for control signal application, and target sensing delays for pinned agents were each randomly sampled for every agent from $U\,(0.001, 0.02)$ s. Control gains are $k_1 = k_2 = 10$, $k_3 = 25$, and $k_4 = 50$. The total simulation duration was 30 s. Fig. 4-1 depicts the tracking error norms and control effort norms, demonstrating robust target tracking despite the uncertainties, noise, and varied delays, over a subset of the 30-second simulation.

CHAPTER 5 COLLABORATIVE STATE ESTIMATION AND TRACKING UNDER DIRECTED PARTIAL RELATIVE STATE FEEDBACK USING DEEP RESIDUAL NEURAL NETWORKS

Obtaining global state information is often impractical in field robotics, motivating relative-only sensing and velocity-free strategies for multi-agent target tracking; yet most adaptive observer-controller designs assume undirected communication and struggle with directed, rooted graphs where asymmetric Laplacians complicate ensemble-level stability and can force multi-hop exchanges. Addressing these limitations, the study considers second-order agents that receive only relative positions and partial relative target states over rooted digraphs, and develops a framework that unifies partial-feedback estimation, relative-only measurements, and online adaptive learning. Core elements include a generalized trackability condition tailored to asymmetric topologies, a locally implementable adaptation law that avoids Laplacian symmetry and multi-hop requirements, and a condensation-graph analysis that exposes how strongly connected components and directed cycles impact trackability under partial information. A deep residual network-based observer learns unknown target dynamics online, while a distributed controller drives agent errors, with a Lyapunov analysis furnishing explicit conditions for exponential convergence to a quantifiable neighborhood of the target. Multi-spacecraft simulations under representative directed topologies corroborate the analysis, with all agents converging within 5 s to a velocity-estimation error of approximately 10.5 m/s and a steady-state tracking error near 1.0 m. The resulting approach removes full-state observability and symmetry assumptions, enabling practical deployment in realistic directed networks.

5.1 Problem Formulation

5.1.1 System Dynamics

Consider a multi-agent system composed of N agents indexed by $i \in \mathcal{V}$ and a target agent indexed by $\{0\}$, with $\overline{\mathcal{V}} \triangleq \mathcal{V} \cup \{0\}$. The dynamics for agent $i \in \mathcal{V}$ are given by

$$\ddot{q}_i = f_i\left(q_i, \dot{q}_i\right) + g_i\left(q_i, \dot{q}_i, t\right) u_i\left(t\right) + \omega_i\left(t\right), \tag{5-1}$$

where $t_0 \in \mathbb{R}_{\geq 0}$ denotes the initial time, $t \geq t_0$ denotes the current time, $q_i : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the agents' unknown generalized position, $\dot{q}_i : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the agents' unknown generalized velocity, the unknown functions $f_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\omega_i : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ represent drift dynamics and exogenous disturbances, respectively, $g_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^{n \times s_i}$ denotes a known control effectiveness matrix, $u_i : \mathbb{R}_{\geq 0} \to \mathbb{R}^{s_i}$ denotes the control input, $s_i \in \mathbb{Z}_{>0}$ denotes the number of control channels, and $\ddot{q}_i : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the agents' unknown generalized acceleration.

The dynamics for the target agent are given by

$$\ddot{q}_0 = f_0 \left(q_0, \dot{q}_0 \right), \tag{5-2}$$

where $q_0 : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the target agent's unknown generalized position, $\dot{q}_0 : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the target agent's unknown generalized velocity, the function $f_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is unknown, and $\ddot{q}_0 : \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ denotes the target agent's unknown generalized acceleration.

Assumptions regarding the dynamical system outlined by (5-1) and (5-2) are provided as follows.

Assumption 5.1. *There exist known constants $\overline{q}_0, \overline{\dot{q}}_0 \in \mathbb{R}_{>0}$ such that $\| q_0 (t) \| \leq \overline{q}_0$ and $\| \dot{q}_0 (t) \| \leq \overline{\dot{q}}_0$ for all $t \in \mathbb{R}_{\geq t_0}$.*

Assumption 5.2. *The functions f_k are globally Lipschitz for all $k \in \overline{\mathcal{V}}$.*

Assumption 5.3. *The functions g_i are full row rank for all $(q_i, \dot{q}_i, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0}$ and are of class C^0. Furthermore, for each fixed $(q_i, \dot{q}_i)$, the map $t \mapsto g_i (q_i, \dot{q}_i, t)$ is uniformly bounded in $t \geq t_0$ for all $i \in \mathcal{V}$.*

Assumption 5.4. *The functions ω_i are locally Lipschitz, and there exist known constants $\overline{\omega}_i \in \mathbb{R}_{>0}$ such that $\| \omega_i (t) \| \leq \overline{\omega}_i$ for all $t \in \mathbb{R}_{\geq t_0}$ and for all $i \in \mathcal{V}$.*

By [74], Assumption 5.3 ensures that the pseudoinverses $g_i^+ : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq t_0} \to \mathbb{R}^{s_i \times n}$ act as right inverses for all $i \in \mathcal{V}$ and $t \in \mathbb{R}_{\geq t_0}$. In addition, the following condition is imposed.

Assumption 5.5. *The functions g_i^+ are of class C^0, and for each fixed $(q_i, \dot{q}_i)$ the map $t \mapsto g_i^+(q_i, \dot{q}_i, t)$ is uniformly bounded in $t \geq t_0$ for all $i \in \mathcal{V}$.*

5.1.2 Communication Topology

Information exchange among the N agents is represented by the static, weighted digraph $G = (\mathcal{V}, E, w)$. The relation $i \to j$ indicates $(i, j) \in E$ and corresponds to agent i having access to information from agent j. For each $i \in \mathcal{V}$, the in-neighborhood is $\mathcal{N}_i \triangleq \{j \in \mathcal{V} : (j, i) \in E, j \neq i\}$.

To model access to the target agent, introduce an external vertex 0 and consider the extended vertex set $\overline{\mathcal{V}} \triangleq \mathcal{V} \cup \{0\}$ together with extended arc set $\overline{E} \subseteq \overline{\mathcal{V}} \times \overline{\mathcal{V}}$. The notation $0 \to i$ indicates $(0, i) \in \overline{E}$; in this case, vertex i is said to be pinned. The extended in-neighborhood is $\overline{\mathcal{N}_i} \triangleq \left\{j \in \overline{\mathcal{V}} : (j, i) \in \overline{E}, j \neq i\right\}$.

5.1.3 Control Objective

Each agent $i \in \mathcal{V}$ can measure the relative position $d_{i,j} \in \mathbb{R}^n$ between itself and its in-neighbors $j \in \mathcal{N}_i$, defined as

$$d_{i,j} \triangleq q_j - q_i. \tag{5-3}$$

Each agent $i \in \mathcal{V}$ can measure its partial relative position with respect to the target agent, defined as

$$y_i \triangleq b_i C_i (q_0 - q_i), \tag{5-4}$$

where $y_i \in \mathbb{R}^{m_i}$, $C_i \in \mathbb{R}^{m_i \times n}$ represents agent i's output matrix (known only to agent i) characterizing its heterogeneous sensing capabilities, $m_i \in \mathbb{Z}_{>0}$ denotes the number of sensing channels, and $b_i \in \{0, 1\}$ denotes a binary indicator of agent i's ability to sense the target.

The primary objective is to design a distributed controller for each agent $i \in \mathcal{V}$ that guides the agent towards the target agent using only the partial relative measurement model. Since relative velocity measurements are unavailable, a secondary objective is to develop a decentralized observer that can estimate the relative velocities using only locally available information from each agent. Additionally, because the target agent's state is unknown, a tertiary objective is to

design a distributed system approximator that reconstructs the target agent's unknown state while simultaneously using online learning techniques to approximate its dynamics.

To quantify the tracking objective, a position tracking error $e_i \in \mathbb{R}^n$ is defined as

$$e_i \triangleq q_0 - q_i, \tag{5-5}$$

for all $i \in \mathcal{V}$. Furthermore, a relative position error $\eta_i \in \mathbb{R}^n$ is defined as $\eta_i \triangleq \sum_{j\in\mathcal{N}_i} A_{ij} d_{i,j} + C_i^\top y_i$, and a relative velocity error is denoted by $\zeta_i \triangleq \dot{\eta}_i$, for all $i \in \mathcal{V}$.

Using (5-3) and (5-5), η_i is expressed in an equivalent analytical form as

$$\eta_i \triangleq \left(b_i C_i^\top C_i e_i - \sum_{j\in\mathcal{N}_i} A_{ij} \left(e_j - e_i \right) \right), \tag{5-6}$$

for all $i \in \mathcal{V}$.

Remark 5.1. The summation $\sum_{j\in\mathcal{N}_i} A_{ij} \left(e_j - e_i \right)$ penalizes state dissimilarity among neighboring agents, while the proportional term $b_i C_i^\top C_i e_i$ penalizes deviations between a pinned agent and the target. Emphasis on inter-agent consensus or target tracking is set by assigning $A_{ij} = k_a$ and $b_i = k_b$, with $k_a, k_b \in \mathbb{R}_{>0}$. If an application requires more rapid consensus among controlled agents, selecting $k_a \gg k_b$ is suitable. Conversely, if rapid convergence to the target is paramount and inter-agent similarity is less critical, then selecting $k_b \gg k_a$ is appropriate.

5.2 Control Design

To facilitate the control design, the filtered tracking error $r_i \in \mathbb{R}^n$ is defined as

$$r_i \triangleq \dot{e}_i + k_1 e_i, \tag{5-7}$$

for all $i \in \mathcal{V}$, where $k_1 \in \mathbb{R}_{>0}$ is a user-defined constant. Denote by $\hat{\eta}_i \in \mathbb{R}^n$ and $\hat{\zeta}_i \in \mathbb{R}^n$ the relative position error and relative velocity error estimates, respectively. The corresponding relative position estimation error $\tilde{\eta}_i \in \mathbb{R}^n$ and relative velocity estimation error $\tilde{\zeta}_i \in \mathbb{R}^n$ are defined

as

$$\tilde{\eta}_i \triangleq \eta_i - \hat{\eta}_i, \quad \tilde{\zeta}_i \triangleq \zeta_i - \hat{\zeta}_i, \tag{5-8}$$

respectively, for all $i \in \mathcal{V}$. Taking the second time-derivative of (5-5), substituting (5-1) and (5-2) into the resulting expression, and adding and subtracting by $f_0\left(\eta_i, \hat{\zeta}_i\right) - f_i\left(\eta_i, \hat{\zeta}_i\right)$ yields

$$\ddot{e}_i = f_0\left(\eta_i, \hat{\zeta}_i\right) - f_i\left(\eta_i, \hat{\zeta}_i\right) - g_i\left(q_i, \dot{q}_i, t\right) u_i\left(t\right) - \omega_i\left(t\right) + \tilde{f}_i\left(q_0, \dot{q}_0, q_i, \dot{q}_i, \eta_i, \hat{\zeta}_i\right), \tag{5-9}$$

where $\tilde{f}_i\left(q_0, \dot{q}_0, q_i, \dot{q}_i, \eta_i, \hat{\zeta}_i\right) \triangleq f_0\left(q_0, \dot{q}_0\right) - f_0\left(\eta_i, \hat{\zeta}_i\right) + f_i\left(\eta_i, \hat{\zeta}_i\right) - f_i\left(q_i, \dot{q}_i\right)$. Substituting (5-9) into the time derivative of (5-7) yields

$$\dot{r}_i = h_i\left(\eta_i, \hat{\zeta}_i\right) - g_i\left(q_i, \dot{q}_i, t\right) u_i\left(t\right) - \omega_i\left(t\right) + \tilde{f}_i\left(q_0, \dot{q}_0, q_i, \dot{q}_i, \eta_i, \hat{\zeta}_i\right) + k_1 \dot{e}_i, \tag{5-10}$$

where $h_i\left(\eta_i, \hat{\zeta}_i\right) \triangleq f_0\left(\eta_i, \hat{\zeta}_i\right) - f_i\left(\eta_i, \hat{\zeta}_i\right) \in \mathbb{R}^n$ for all $i \in \mathcal{V}$.

5.2.1 Residual Neural Network Function Approximation

Physics-based models often struggle to represent the nonlinear, coupled dynamics present in multi-agent tracking problems because they rely on simplifying assumptions and uncertain parameters. ResNets provide a data-driven alternative, enabling the unknown dynamics to be learned without explicit knowledge of the governing physics. Accordingly, a ResNet function approximator is adopted to obtain more accurate and robust models of nonlinear dynamics.

To approximate the unknown dynamics $h_i\left(\eta_i, \hat{\zeta}_i\right)$ in (5-10) for all $i \in \mathcal{V}$, define the ResNets' input vector $\kappa_i \in \mathbb{R}^{2n}$ as $\kappa_i \triangleq \begin{bmatrix} \eta_i^\top & \hat{\zeta}_i^\top \end{bmatrix}^\top \in \Omega$, where $\Omega \subset \mathbb{R}^{2n}$ is a compact set over which the universal approximation property holds for all $i \in \mathcal{V}$. The ResNet-based approximation of $h_i\left(\eta_i, \hat{\zeta}_i\right)$ is given by $\Psi_i\left(\kappa_i, \hat{\theta}_i\right)$, where $\Psi_i : \mathbb{R}^{2n} \times \mathbb{R}^{p_i} \to \mathbb{R}^n$ denotes the ResNet architecture mapping and $\hat{\theta}_i \in \mathbb{R}^{p_i}$ denotes the adaptive parameter estimates, for all $i \in \mathcal{V}$. By Assumption 5.2, each h_i is continuous on Ω, for all $i \in \mathcal{V}$.

Fix a user-prescribed accuracy $\overline{\varepsilon} > 0$. By the universal approximation property on the compact domain Ω (e.g., [94, Corollary 5.2]), there exist corresponding ResNet architectures (hence $p_i \in \mathbb{Z}_{>0}$) and compact parameter search spaces $\mho_i \subset \mathbb{R}^{p_i}$ with C^∞ boundaries and $\mathbf{0}_{p_i} \in \operatorname{int}(\mho_i)$ such that $\inf_{\vartheta \in \mho_i} \sup_{\kappa_i \in \Omega} \|h_i(\kappa_i) - \Psi_i(\kappa_i, \vartheta)\| \leq \overline{\varepsilon}$, for all $i \in \mathcal{V}$.

For subsequent analysis, introduce $\theta_i^* \in \arg\min_{\vartheta \in \mho_i} \sup_{\kappa_i \in \Omega} \|h_i(\kappa_i) - \Psi_i(\kappa_i, \vartheta)\|$ under these fixed $\Psi_i(\cdot, \cdot)$, for all $i \in \mathcal{V}$. By joint continuity of $(\kappa_i, \vartheta) \mapsto \|h_i(\kappa_i) - \Psi_i(\kappa_i, \vartheta)\|$ on $\Omega \times \mho_i$ and compactness of Ω, the map $\vartheta \mapsto \sup_{\kappa_i \in \Omega} \|h_i(\kappa_i) - \Psi_i(\kappa_i, \vartheta)\|$ is continuous for all $i \in \mathcal{V}$. Since each $\mho_i$ is compact, a minimizer exists by Weierstrass's Theorem (e.g., [36, Theorem 10.55]). Fix such a θ_i^*, for all $i \in \mathcal{V}$. With this choice, $\sup_{\kappa_i \in \Omega} \left\| h_i(\kappa_i) - \Psi_i\left(\kappa_i, \theta_i^*\right) \right\| \leq \overline{\varepsilon}$, for all $i \in \mathcal{V}$.

Using these fixed parameters, write

$$h_i(\kappa_i) = \Psi_i\left(\kappa_i, \theta_i^*\right) + \varepsilon_i(\kappa_i), \tag{5-11}$$

where $\varepsilon_i(\kappa_i) \triangleq h_i(\kappa_i) - \Psi_i\left(\kappa_i, \theta_i^*\right)$ satisfies the uniform bound $\sup_{\kappa_i \in \Omega} \|\varepsilon_i(\kappa_i)\| \leq \overline{\varepsilon}$ for all $i \in \mathcal{V}$.

5.2.2 Distributed Observer-based Control Design

The controller, filter, observer, and adaptation law are constructed to ensure closed-loop stability of the system with dynamics defined in (5-1) and (5-2). The design cancels the cross-coupled terms in the Lyapunov derivative while bounding or exploiting the remaining terms to guarantee negative definiteness. Auxiliary variables such as a filtered estimation error and tunable gains are introduced to realize this objective. The subsequent stability analysis establishes boundedness and convergence of the closed-loop signals and justifies the design of the control input, observer, and adaptation law.

To facilitate the distributed observer design, a filtered estimation error $\tilde{r}_i \in \mathbb{R}^n$ is defined as

$$\tilde{r}_i \triangleq \dot{\tilde{\eta}}_i + k_3 \tilde{\eta}_i + \rho_i, \tag{5-12}$$

for all $i \in \mathcal{V}$, where $k_3 \in \mathbb{R}_{>0}$ is a user-defined constant, and $\rho_i \in \mathbb{R}^n$ is designed as

$$\rho_i(t) = -(k_3+k_4)\tilde{\eta}_i(t) + \left(1 - k_3^2 - k_3 k_4\right)\int_{t_0}^{t} \tilde{\eta}_i(\tau)\,\mathrm{d}\tau - (k_3+k_4+k_5)\int_{t_0}^{t} \rho_i(\tau)\,\mathrm{d}\tau,$$
$$\rho_i(0) = \mathbf{0}_n, \tag{5-13}$$

where $k_4, k_5 \in \mathbb{R}_{>0}$ are user-defined constants for all $i \in \mathcal{V}$. The distributed observer is designed as

$$\begin{aligned}
\dot{\hat{\eta}}_i &= \hat{\zeta}_i, \\
\dot{\hat{\zeta}}_i &= \sum_{j\in\mathcal{N}_i} A_{ij}\left(g_j u_j - g_i u_i\right) - b_i C_i^\top C_i g_i u_i + b_i C_i^\top C_i \Psi_i\left(\kappa_i, \hat{\theta}_i\right) \\
&\quad - \sum_{j\in\mathcal{N}_i} A_{ij}\left(\Psi_j\left(\kappa_j, \hat{\theta}_j\right) - \Psi_i\left(\kappa_i, \hat{\theta}_i\right)\right) - \left(k_3^2 - 2\right)\tilde{\eta}_i - (2k_3 + k_4 + k_5)\rho_i,
\end{aligned} \tag{5-14}$$

for all $i \in \mathcal{V}$. Based on the subsequent stability analysis, the control input is designed as

$$u_i = g_i^+\left(\Psi_i\left(\kappa_i, \hat{\theta}_i\right) + k_2\left(k_1\eta_i + \hat{\zeta}_i - k_3\tilde{\eta}_i - \rho_i\right)\right) - b_i\left(k_1^2 + 1\right) g_i^+ C_i^\top y_i, \tag{5-15}$$

where $k_2 \in \mathbb{R}_{>0}$ is a user-defined constant. Similarly, the adaptation law for the ResNet is designed as

$$\dot{\hat{\theta}}_i = \mathrm{proj}\left(\Theta_i\right), \tag{5-16}$$

where

$$\begin{aligned}
\Theta_i \triangleq\ & k_1\Gamma_i\left(\frac{\partial\Psi_i\left(\kappa_i, \hat{\theta}_i\right)}{\partial\hat{\theta}_i}\right)^\top C_i^\top y_i + |\mathcal{N}_i|\,\Gamma_i\left(\frac{\partial\Psi_i\left(\kappa_i, \hat{\theta}_i\right)}{\partial\hat{\theta}_i}\right)^\top \left(k_3\tilde{\eta}_i + \rho_i - \hat{\zeta}_i\right) \\
& + b_i\Gamma_i\left(\frac{\partial\Psi_i\left(\kappa_i, \hat{\theta}_i\right)}{\partial\hat{\theta}_i}\right)^\top C_i^\top C_i\left(k_3\tilde{\eta}_i + \rho_i - \hat{\zeta}_i\right) - k_6\Gamma_i\left(\xi_i\hat{\theta}_i - \sum_{j\in\mathcal{N}_i} A_{ij}\left(\hat{\theta}_j - \hat{\theta}_i\right)\right),
\end{aligned}$$

where $k_6 \in \mathbb{R}_{>0}$ is a user-defined forgetting rate, $\xi_i \in \mathbb{R}_{>0}$ are user-defined nodal regularization gains, and $\Gamma_i \in \mathbb{R}^{p_i\times p_i}$ are user-defined symmetric positive-definite learning rate matrices, for all $i \in \mathcal{V}$. Here, the projection operator is defined as in [52, Appendix E, Equation E.4] and ensures the parameter estimates satisfy $\left\|\hat{\theta}_i(t)\right\| \le \overline{\theta}$ for all $t \in \mathbb{R}_{\ge t_0}$ and $i \in \mathcal{V}$, where $\overline{\theta} \triangleq \sqrt{\theta_{\max}^2 + \overline{\epsilon}}$,

$\theta_{\max} \triangleq \max_{i\in\mathcal{V}} \sup_{\vartheta\in\mathfrak{V}_i} \|\vartheta\|$, and $\overline{\epsilon} > 0$ is a margin introduced to ensure Lipschitz continuity of the projection mapping.

Remark 5.2. The term $-k_6\left(\xi_i\hat{\theta}_i - \sum_{j\in\mathcal{N}_i} A_{ij}\left(\hat{\theta}_j - \hat{\theta}_i\right)\right)$ introduces a distributed parameter coupling mechanism. This component can be interpreted as a consensus-based regularization, which adjusts the local parameter $\hat{\theta}_i$ by incorporating information from the neighborhood $\mathcal{N}_i$ in the communication graph. This coupling biases the parameter update towards consistency across neighboring agents, promoting consensus in the estimated weights. From the perspective of distributed optimization, this term is analogous to the strategy of minimizing a global cost function through local interactions, a characteristic feature of distributed gradient-based algorithms (e.g., [92]). The interplay between the gradient-like term and the consensus protocol enables agents to refine their estimates using both local data and information propagated through the network, which is critical for achieving distributed adaptation.

5.2.3 Ensemble Representation

To aid in the stability analysis, the interaction matrix $\mathcal{H} \in \mathbb{R}^{nN\times nN}$ is defined as

$$\mathcal{H} \triangleq (L \otimes I_n) + C, \tag{5-17}$$

where $C \triangleq \text{blkdiag}\left\{b_1 C_1^\top C_1, \ldots, b_N C_N^\top C_N\right\} \in \mathbb{R}^{nN\times nN}$.

Lemma 5.1. *The interaction matrix $\mathcal{H}$, as defined in (5-17), is a linear operator acting on the vector space $\mathbb{R}^{nN}$. More precisely, for any vector $\mathtt{x} \in \mathbb{R}^{nN}$ partitioned into components $\mathtt{x}_i \in \mathbb{R}^n$ such that $\mathtt{x} = [\mathtt{x}_i]_{i\in[N]} \in \mathbb{R}^{nN}$, the action of $\mathcal{H}$ on $\mathtt{x}$ yields a vector $\mathtt{y} = \mathcal{H}\mathtt{x} \in \mathbb{R}^{nN}$, similarly partitioned as $\mathtt{y} = [\mathtt{y}_i]_{i\in[N]} \in \mathbb{R}^{nN}$, where each $\mathtt{y}_i \in \mathbb{R}^n$ is given by*
$\mathtt{y}_i = b_i C_i^\top C_i \mathtt{x}_i - \sum_{j\in\mathcal{N}_i} A_{ij}\left(\mathtt{x}_j - \mathtt{x}_i\right)$, *for $i \in [N]$.*

Proof. For any $A \in \mathbb{R}^{N\times N}$ and $\mathtt{z} = [\mathtt{z}_i]_{i\in[N]} \in \mathbb{R}^{nN}$ with $\mathtt{z}_i \in \mathbb{R}^n$, $(A \otimes I_n)\,\mathtt{z} = \left[\sum_{j=1}^N A_{ij}\mathtt{z}_j\right]_{i\in[N]}$. Following the structure of L, $L_{ii} = \sum_{j\in\mathcal{N}_i} A_{ij}$ and $L_{ij} = -A_{ij}$ for $j \in \mathcal{N}_i$ and 0 otherwise. Hence, $\sum_{j=1}^N L_{ij}\mathtt{x}_j = \left(\sum_{j\in\mathcal{N}_i} A_{ij}\right)\mathtt{x}_i - \sum_{j\in\mathcal{N}_i} A_{ij}\mathtt{x}_j = \sum_{j\in\mathcal{N}_i} A_{ij}\left(\mathtt{x}_i - \mathtt{x}_j\right)$. Since C is block diagonal, $(C\mathtt{x})_i = b_i C_i^\top C_i \mathtt{x}_i$. Thus, $\mathtt{y}_i = ((L \otimes I_n)\,\mathtt{x})_i + (C\mathtt{x})_i = b_i C_i^\top C_i \mathtt{x}_i - \sum_{j\in\mathcal{N}_i} A_{ij}\left(\mathtt{x}_j - \mathtt{x}_i\right)$. □

Using Lemma 5.1, (5-6) is expressed in an ensemble form as

$$\eta = \mathcal{H} e, \tag{5-18}$$

where $\eta \triangleq [\eta_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$ and $e \triangleq [e_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Using (5-7) and (5-18) yields the expression

$$\mathcal{H} r = \dot{\eta} + k_1 \eta, \tag{5-19}$$

where $r \triangleq [r_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Using (5-4), (5-18) and (5-19), (5-15) is expressed in an ensemble form as

$$u = g^{+} \left(\widehat{\Psi} + k_2 \mathcal{H} r - k_2 \tilde{r} - \left(k_1^2 + 1 \right) C e \right), \tag{5-20}$$

where $u \triangleq [u_i]_{i \in \mathcal{V}} \in \mathbb{R}^{\sum_{i \in \mathcal{V}} s_i}$, $g^{+} \triangleq \text{blkdiag}\left\{ g_1^{+}, \ldots g_N^{+} \right\} \in \mathbb{R}^{\sum_{i \in \mathcal{V}} s_i \times nN}$, $\widehat{\Psi} \triangleq \left[\Psi_i \left(\kappa_i, \hat{\theta}_i \right) \right]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$, $\tilde{r} \triangleq [\tilde{r}_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Using (5-17), (5-14) is expressed in an ensemble form as

$$\begin{aligned} \dot{\hat{\eta}} &= \hat{\zeta}, \\ \dot{\hat{\zeta}} &= \mathcal{H} \left(\widehat{\Psi} - g u \right) - \left(k_3^2 - 2 \right) \tilde{\eta} - \left(2 k_3 + k_4 + k_5 \right) \rho, \end{aligned} \tag{5-21}$$

where $\hat{\eta} \triangleq [\hat{\eta}_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$, $\hat{\zeta} \triangleq \left[\hat{\zeta}_i \right]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$, $g \triangleq \text{blkdiag}\left\{ g_1, \ldots g_N \right\} \in \mathbb{R}^{nN \times \sum_{i \in \mathcal{V}} s_i}$, $\tilde{\eta} \triangleq [\tilde{\eta}_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$, and $\rho \triangleq [\rho_i]_{i \in \mathcal{V}} \in \mathbb{R}^{nN}$. Furthermore, the time-derivative of (5-13) is expressed in an ensemble form as

$$\dot{\rho} \triangleq \tilde{\eta} - \left(k_3 + k_4 \right) \tilde{r} - k_5 \rho. \tag{5-22}$$

Substituting (5-20) and (5-11) into the ensemble representation of (5-10) and using the ensemble representation of (5-7) yields

$$\dot{r} = \Psi^{*} - \widehat{\Psi} - k_2 \mathcal{H} r + k_2 \tilde{r} + k_1 r + \left(\left(k_1^2 + 1 \right) C - k_1^2 I_{nN} \right) e - \omega + \tilde{f} + \varepsilon, \tag{5-23}$$

where $\Psi^* \triangleq \left[\Psi_i\left(\kappa_i, \theta_i^*\right)\right]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $\omega \triangleq [\omega_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, $\tilde{f} \triangleq \left[\tilde{f}_i\right]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, and $\varepsilon \triangleq [\varepsilon_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$.

Taking the time-derivative of (5-12), using (5-9), (5-11), (5-12), and (5-18), and then substituting (5-21) and (5-22) into the ensemble representation of the resulting expression yields

$$\dot{\tilde{r}} = \mathcal{H}\left(\Psi^* - \widehat{\Psi} + \varepsilon - \omega + \tilde{f}\right) - \tilde{\eta} + (k_3 + k_4)\,\rho - k_4\tilde{r}, \tag{5-24}$$

where $\tilde{r} \triangleq [\tilde{r}_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$.

5.3 Stability Analysis

The ResNet described in (5-11) is nonlinear with respect to its parameters. Designing adaptive controllers and performing stability analyses for nonlinearly parameterized systems present significant theoretical challenges. To address this nonlinearity, a first-order application of Taylor's theorem about the current parameter estimate is used. Define the parameter estimation error $\tilde{\theta}_i \in \mathbb{R}^{p_i}$ as

$$\tilde{\theta}_i = \theta_i^* - \hat{\theta}_i, \tag{5-25}$$

for all $i \in \mathcal{V}$. Applying first-order Taylor's theorem to $\Psi_i\left(\kappa_i, \theta_i^*\right)$ about the point $\left(\kappa_i, \hat{\theta}_i\right)$ and using (5-25) yields

$$\Psi_i\left(\kappa_i, \theta_i^*\right) = \Psi_i\left(\kappa_i, \hat{\theta}_i\right) + \frac{\partial}{\partial \hat{\theta}_i}\Psi_i\left(\kappa_i, \hat{\theta}_i\right)\tilde{\theta}_i + R_i\left(\kappa_i, \tilde{\theta}_i\right), \tag{5-26}$$

where $\frac{\partial}{\partial \hat{\theta}_i}\Psi_i\left(\kappa_i, \hat{\theta}_i\right) \triangleq \left.\frac{\partial}{\partial \vartheta_i}\right|_{\hat{\theta}_i} \Psi_i\left(\kappa_i, \vartheta_i\right) \in \mathbb{R}^{n\times p_i}$ and $R_i : \mathbb{R}^{2n} \times \mathbb{R}^{p_i} \to \mathbb{R}^n$ denotes the first Lagrange remainder for all $i \in \mathcal{V}$. Substituting (5-26) into (5-23) yields

$$\dot{r} = \frac{\partial \widehat{\Psi}}{\partial \hat{\theta}}\tilde{\theta} - k_2\mathcal{H}r + k_2\tilde{r} + k_1 r - \omega + \tilde{f} + R + \varepsilon + \left(\left(k_1^2 + 1\right)C - k_1^2 I_{nN}\right)e, \tag{5-27}$$

where $\hat{\theta} \triangleq \left[\hat{\theta}_i\right]_{i\in\mathcal{V}} \in \mathbb{R}^{pN}$, $\frac{\partial \widehat{\Psi}}{\partial \hat{\theta}} \triangleq \text{blkdiag}\left\{\frac{\partial}{\partial \hat{\theta}_1}\Psi_1\left(\kappa_1, \hat{\theta}_1\right), \ldots, \frac{\partial}{\partial \hat{\theta}_N}\Psi_N\left(\kappa_N, \hat{\theta}_N\right)\right\} \in \mathbb{R}^{nN\times pN}$, $\tilde{\theta} \triangleq \left[\tilde{\theta}_i\right]_{i\in\mathcal{V}} \in \mathbb{R}^{pN}$, and $R\left(\kappa, \tilde{\theta}\right) \triangleq \left[R_i\left(\kappa_i, \tilde{\theta}_i\right)\right]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, with $p \triangleq \sum_{i\in\mathcal{V}} p_i$ and

$\kappa \triangleq [\kappa_i]_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$. Substituting (5-26) into (5-24) yields

$$\dot{\tilde{r}} = \mathcal{H}\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\tilde{\theta} + R + \varepsilon - \omega + \tilde{f}\right) + (k_3 + k_4)\,\rho - \tilde{\eta} - k_4\tilde{r}. \tag{5-28}$$

The ensemble representation of (5-16) is expressed as $\dot{\hat{\theta}} \in \mathbb{R}^{pN}$, where

$$\dot{\hat{\theta}} = \begin{bmatrix} \text{proj}\,(\Theta_1)^\top & \cdots & \text{proj}\,(\Theta_N)^\top \end{bmatrix}^\top. \tag{5-29}$$

Substituting (5-29) into the ensemble representation of (5-25) yields

$$\dot{\tilde{\theta}} = -\begin{bmatrix} \text{proj}\,(\Theta_1)^\top & \cdots & \text{proj}\,(\Theta_N)^\top \end{bmatrix}^\top. \tag{5-30}$$

5.3.1 Trackability

This work examines scenarios in which multiple agents with limited sensing capabilities collectively track a single target. In this setting, each agent typically measures only a subset of the target's states. Without sharing these measurements, the group may be unable to reconstruct the target's full state. For instance, in a two-dimensional setting, two agents that each measure only distinct coordinates of the target's position and do not communicate can at best align themselves collinearly with the target. By contrast, if they exchange their partial measurements, they can collaborate to reconstruct the complete state, enabling effective multi-agent tracking. The trackability condition formalizes this principle by requiring that the collective sensing capabilities of all agents provide sufficient information for accurate state estimation. In particular, the trackability condition enables the feasibility of the tracking objective and, together with the positivity of the interaction term introduced below (obtained either from a local feasibility condition or from a balanced or undirected communication topology), yields strict positivity of the interaction matrix used in the subsequent stability analysis.

Recall that communication among agents is modeled by a static weighted digraph $G = (\mathcal{V}, E, w)$. Information must flow from the sensing agents to the rest of the network, which is captured by the following pinning condition at the root.

Assumption 5.6. *The digraph G is rooted, and at least one root vertex of G senses the target 0; that is, there exists a root vertex $r \in \mathcal{V}$ with $0 \to r$. Consequently, there exists at least one $i \in \mathcal{V}$ with $b_i > 0$, where $b_i \in \{0, 1\}$.*

For general unbalanced digraphs, the next local condition ensures strict positivity of the quadratic form of $\mathcal{H}$.

Assumption 5.7. *Each agent $i \in \mathcal{V}$ satisfies $\lambda_{\min}\left(b_i C_i^\top C_i\right) + \frac{1}{2}\left(d_i^{in} - d_i^{out}\right) > 0$.*

The aggregate sensing required in the root components to ensure feasibility of the tracking problem is as follows.

Definition 5.1. (Trackability) The target is trackable if there exists a $\mu > 0$ such that

$$\sum_{k\in\mathcal{S}} b_k C_k^\top C_k \succeq \mu I_n \iff \lambda_{\min}\left(\sum_{k\in\mathcal{S}} b_k C_k^\top C_k\right) \geq \mu.$$

Equivalently, the target is trackable if rank $\left(\sum_{k\in\mathcal{S}} b_k C_k^\top C_k\right) = n$. Here, $\mathcal{S}$ is the unique root SCC.

The next definition isolates the minimal algebraic requirement under which the target state can be uniquely reconstructed from the ensemble of partial measurements and the known agent states at a given time.

Definition 5.2. (Target Identifiability) The target state $q_0 \in \mathbb{R}^n$ is said to be identifiable from $[\{q_i, y_i\}]_{i\in\mathcal{V}}$ if the measurement map $\Upsilon : \mathbb{R}^n \to \mathbb{R}^{\sum_{i=1}^N m_i}$, $\Upsilon(q) \triangleq [b_i C_i (q - q_i)]_{i\in\mathcal{V}}$ is injective, i.e., $\Upsilon(q)$ is full column rank.

Two auxiliary results are needed.

Lemma 5.2. *Let $A \in \mathbb{R}^{n\times n}$ satisfy $x^\top A x > 0$ for every $x \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$. Then every eigenvalue λ of A satisfies Re $(\lambda) > 0$.*

Proof. Decompose $A = A_{\text{Sym}} + A_{\text{Skew}}$, where $A_{\text{Sym}} \triangleq \frac{1}{2}(A + A^\top)$ and $A_{\text{Skew}} \triangleq \frac{1}{2}(A - A^\top)$ (e.g., [21, Section 0.2.5]). For any $x \in \mathbb{R}^n$,
$x^\top A_{\text{Skew}} x = \frac{1}{2} x^\top (A - A^\top) x = \frac{1}{2} x^\top (A - A^\top)^\top x = -\frac{1}{2} x^\top (A - A^\top) x = -x^\top A_{\text{Skew}} x = 0$, so
$x^\top A x = x^\top A_{\text{Sym}} x$. Thus, A_{Sym} is positive definite. If $Av = \lambda v$ with $v \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}$, then
$\text{Re}(\lambda)\, v^* v = \text{Re}(v^* A v) = \text{Re}\left(v^* A_{\text{Sym}} v\right) + \text{Re}(v^* A_{\text{Skew}} v) = v^* A_{\text{Sym}} v > 0$ because $v^* A_{\text{Sym}} v$ is real and positive and $v^* A_{\text{Skew}} v$ is purely imaginary for real skew-symmetric A_{Skew}. Hence,
$\text{Re}(\lambda) > 0$. □

Lemma 5.3. *Let $G = (\mathcal{V}, E, w)$ be a rooted digraph with N vertices and in-Laplacian $L = D - A$, where D is the in-degree matrix and A is the adjacency matrix. Then* $\dim(\ker L) = 1$ *and* $\ker L = \text{span}\{\mathbf{1}_N\}$.

Proof. Since G is rooted, there exists $r \in \mathcal{V}$ with the property that for every $j \in \mathcal{V}$, the relation $r \rightsquigarrow j$ holds. Hence, the reachable set of r satisfies $\mathcal{R}(r) = \{i \in \mathcal{V} : r \rightsquigarrow i\} \cup \{r\} = \mathcal{V}$. Thus, $\mathcal{V}$ itself is a reach, and by maximality no other distinct reach can occur. Therefore, the number of reaches of G is one. Since each row of L sums to zero, $\mathbf{1}_N \in \ker L$.

By [6, Corollary 4.2], the algebraic multiplicity of the eigenvalue 0 is equal to the number of reaches. Since the number of reaches is one and $\mathbf{1}_N \in \ker L$, it follows that $\dim(\ker L) = 1$ and $\ker L = \text{span}\{\mathbf{1}_N\}$. □

The main result connects the definiteness of $\mathcal{H}$, trackability, and target identifiability.

Theorem 5.1. *Suppose G is rooted.*

1. *If Assumption 5.7 holds, then $x^\top \mathcal{H} x > 0$ for all $x \neq \mathbf{0}_{nN}$.*
2. *If $x^\top \mathcal{H} x > 0$ for all $x \neq \mathbf{0}_{nN}$, then the target is trackable.*
3. *Trackability $\implies$ identifiability.*

Consequently, under Assumption 5.7 the target is trackable and identifiable, and every eigenvalue of $\mathcal{H}$ has positive real part.

Proof. First, strict positivity of the quadratic form under Assumption 5.7 is established. Write $x = [x_i]_{i\in[N]} \in \mathbb{R}^{nN}$ with $x_i \in \mathbb{R}^n$ for all $i \in \mathcal{V}$. Using $L = D - A$ and the inequality $2u^\top v \leq \left(\|u\|^2 + \|v\|^2\right)$,

$$\begin{aligned} x^\top (L \otimes I_n)\, x &= \sum_{i=1}^{N} d_i^{\text{in}} \|x_i\|^2 - \sum_{(j,i)\in E} A_{ij} x_i^\top x_j, \\ &\geq \sum_{i=1}^{N} d_i^{\text{in}} \|x_i\|^2 - \frac{1}{2} \sum_{(j,i)\in E} A_{ij} \left(\|x_i\|^2 + \|x_j\|^2\right), \\ &= \frac{1}{2} \sum_{i=1}^{N} \left(d_i^{\text{in}} - d_i^{\text{out}}\right) \|x_i\|^2 . \end{aligned}$$

Moreover,

$$x^\top C x = \sum_{i=1}^{N} x_i^\top \left(b_i C_i^\top C_i\right) x_i \geq \sum_{i=1}^{N} \lambda_{\min}\left(b_i C_i^\top C_i\right) \|x_i\|^2 .$$

Consequently, by Assumption 5.7,

$$\begin{aligned} x^\top \mathcal{H} x &= x^\top (L \otimes I_n)\, x + x^\top C x, \\ &\geq \sum_{i=1}^{N} \left(\lambda_{\min}\left(b_i C_i^\top C_i\right) + \frac{1}{2}\left(d_i^{\text{in}} - d_i^{\text{out}}\right)\right) \|x_i\|^2 , \\ &> 0, \end{aligned}$$

for every $x \neq \mathbf{0}_{nN}$.

Second, strict positivity of the quadratic form implies trackability is shown. The argument proceeds by contrapositive. Suppose the target is not trackable. Then, there exists a unique root SCC $\mathcal{S} \subset \mathcal{V}$ with

$$\operatorname{rank}\left(\sum_{i\in\mathcal{S}} b_i C_i^\top C_i\right) < n.$$

Hence, there exists $k \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$ such that

$$\sum_{i \in \mathcal{S}} b_i \left\| C_i k \right\|^2 = k^\top \left(\sum_{i \in \mathcal{S}} b_i C_i^\top C_i \right) k = 0,$$

which forces $C_i k = \mathbf{0}_{m_i}$ for every $i \in \mathcal{S}$ with $b_i > 0$. Permute vertices so that $\mathcal{S}$ forms the leading block. There exists a permutation matrix $P \in \mathbb{R}^{N \times N}$ such that $\widetilde{L} \triangleq PLP^\top \in \mathbb{R}^{N \times N}$ is block lower triangular, with the Laplacian $L_{\mathcal{S}}$ of $\mathcal{S}$ in the top-left block and with the off-diagonal block $\mathbf{0}_{|\mathcal{S}| \times |\mathcal{V} \setminus \mathcal{S}|}$. Set $\widetilde{P} \triangleq P \otimes I_n \in \mathbb{R}^{nN} \times \mathbb{R}^{nN}$ and

$$\tilde{x} \triangleq \begin{bmatrix} \mathbf{1}_{|\mathcal{S}|} \otimes k \\ \mathbf{0}_{n(N - |\mathcal{S}|)} \end{bmatrix} \neq \mathbf{0}_{nN}.$$

Because each row of $L_{\mathcal{S}}$ sums to zero, $\mathbf{1}_{|\mathcal{S}|}^\top L_{\mathcal{S}} \mathbf{1}_{|\mathcal{S}|} = 0$. Using $(u \otimes v)^\top (A \otimes I_n)(u \otimes v) = (u^\top A u) \left\| v \right\|^2$,

$$\tilde{x}^\top \left(\widetilde{L} \otimes I_n \right) \tilde{x} = \left(\mathbf{1}_{|\mathcal{S}|}^\top L_{\mathcal{S}} \mathbf{1}_{|\mathcal{S}|} \right) \left\| k \right\|^2 = 0.$$

Moreover, $\widetilde{P} C \widetilde{P}^\top \in \mathbb{R}^{nN} \times \mathbb{R}^{nN}$ is block diagonal with diagonal blocks $b_i C_i^\top C_i$ in the permuted order, implying

$$\tilde{x} \widetilde{P} C \widetilde{P}^\top \tilde{x} = \sum_{i \in \mathcal{S}} b_i \left\| C_i k \right\|^2 = 0,$$

because $C_i k = \mathbf{0}_{m_i}$ for all $i \in \mathcal{S}$ with $b_i > 0$. Combining with $\tilde{x}^\top \left(\widetilde{L} \otimes I_n \right) \tilde{x} = 0$ gives $\tilde{x}^\top \left(\left(\widetilde{L} \otimes I_n \right) + \widetilde{P} C \widetilde{P}^\top \right) \tilde{x} = 0$. Recalling the definitions of $\widetilde{L}$ and $\widetilde{P}$ yields the equality $\widetilde{L} \otimes I_n = \widetilde{P} (L \otimes I_n) \widetilde{P}^\top$. Using this expression yields $\tilde{x}^\top \left(\left(\widetilde{L} \otimes I_n \right) + \widetilde{P} C \widetilde{P}^\top \right) \tilde{x} = \left(\widetilde{P}^\top \tilde{x} \right)^\top ((L \otimes I_n) + C) \widetilde{P}^\top \tilde{x}$. Therefore, defining $x \triangleq \widetilde{P}^\top \tilde{x} \in \mathbb{R}^{nN}$ yields $x^\top \mathcal{H} x = 0$. To show that $x \neq \mathbf{0}_{nN}$, recall that since P is a permutation matrix, $P^\top P = I_N$. Therefore $\widetilde{P} = P \otimes I_n$ is also a permutation matrix with $\widetilde{P}^\top \widetilde{P} = I_{nN}$ and $\ker \left(\widetilde{P}^\top \widetilde{P} \right) = \{\mathbf{0}_{nN}\}$, so $\tilde{x} \neq \mathbf{0}_{nN}$ implies $x = \widetilde{P}^\top \tilde{x} \neq \mathbf{0}_{nN}$. Thus, there exists $x \in \mathbb{R}^{nN} \setminus \{\mathbf{0}_{nN}\}$ such that $x^\top \mathcal{H} x = 0$, contradicting strict positivity. Therefore, strict positivity implies trackability.

Third, trackability implying identifiability is shown. Define the stacked sensing matrix $\overline{C}_b \triangleq [b_i C_i]_{i\in\mathcal{V}} \in \mathbb{R}^{\sum_{i=1}^N m_i \times n}$ and the stacked data $\overline{s} \triangleq [y_i + b_i C_i q_i]_{i\in\mathcal{V}} \in \mathbb{R}^{\sum_{i=1}^N m_i}$. The measurement relation $y_i = b_i C_i (q_0 - q_i)$ gives $\overline{s} = \overline{C}_b q_0$. By trackability, for the unique root SCC $\mathcal{S}$, $\sum_{i\in\mathcal{S}} b_i C_i^\top C_i \succeq \mu I_n$ for $\mu > 0$. Summing these inequalities over the root SCC and using $b_i C_i^\top C_i \succeq 0$ outside the root yields

$$\overline{C}_b^\top \overline{C}_b = \sum_{i\in\mathcal{V}} b_i C_i^\top C_i \succeq \mu I_n,$$

because $b_i^2 = b_i$ for all $b_i \in \{0, 1\}$. Thus $\overline{C}_b$ has full column rank, the map $q_0 \mapsto \overline{C}_b q_0$ is injective, and q_0 is uniquely determined from the data by $q_0 = \overline{C}_b^+ \overline{s}$.

Finally, since $x^\top \mathcal{H} x > 0$ for all $x \neq \mathbf{0}_{nN}$ under Assumption 5.7, Lemma 5.2 applies to $\mathcal{H}$ and yields $\mathrm{Re}\,(\lambda) > 0$ for every eigenvalue λ of $\mathcal{H}$. □

Under weight-balanced or undirected topologies, the reverse implication also holds: trackability implies $x^\top \mathcal{H} x > 0$ for all $x \neq \mathbf{0}_{nN}$ and Assumption 5.7 is unnecessary in this case.

Corollary 5.1. *Suppose Assumption 5.6 holds and either G is weight-balanced ($d_i^{in} = d_i^{out}$ for all $i \in \mathcal{V}$) or undirected ($A_{ij} = A_{ji}$ for all $i, j \in \mathcal{V}$). If the target is trackable, then $x^\top \mathcal{H} x > 0$ for all $x \in \mathbb{R}^{nN} \setminus \{\mathbf{0}_{nN}\}$.*

Proof. As before, write $x = [x_i]_{i\in[N]} \in \mathbb{R}^{nN}$ with $x_i \in \mathbb{R}^n$. Using $L = D - A$ and the definition of A,

$$x^\top (L \otimes I_n)\, x = \sum_{(j,i)\in E} A_{ij} \|x_i\|^2 - \sum_{(j,i)\in E} A_{ij} x_i^\top x_j.$$

If G is weight-balanced, then

$\sum_{(j,i)\in E} A_{ij} \|x_i\|^2 = \sum_{(j,i)\in E} A_{ij} \left\|x_j\right\|^2 = \frac{1}{2} \sum_{(j,i)\in E} A_{ij} \left(\|x_i\|^2 + \left\|x_j\right\|^2\right)$. If G is undirected, then A is symmetric and the same identity holds. Hence,

$$x^\top (L \otimes I_n)\, x = \frac{1}{2} \sum_{(j,i)\in E} A_{ij} \left\|x_i - x_j\right\|^2 \geq 0,$$

so $L \otimes I_n \succeq 0$. By Lemma 5.3, $\ker L = \operatorname{span}\{\mathbf{1}_N\}$. Therefore,

$$\ker\left(L \otimes I_n\right) = \ker L \otimes \mathbb{R}^n = \operatorname{span}\{\mathbf{1}_N\} \otimes \mathbb{R}^n.$$

Let $x \in \ker\left(L \otimes I_n\right) \setminus \{\mathbf{0}_{nN}\}$, so $x = \mathbf{1}_N \otimes \mathtt{u}$ with $\mathtt{u} \neq \mathbf{0}_n$. Trackability ensures that there exists some $\mu > 0$ such that $\sum_{i \in \mathcal{V}} C_i^\top C_i \succeq \mu I_n$ and a unique root SCC $\mathcal{S}$ with

$$\begin{aligned} x^\top C x = \sum_{i=1}^{N} \mathtt{u}^\top \left(b_i C_i^\top C_i\right) \mathtt{u} &\geq \mathtt{u}^\top \left(\sum_{i \in \mathcal{S}} b_i C_i^\top C_i\right) \mathtt{u}, \\ &\geq \mu \left\|\mathtt{u}\right\|^2 > 0. \end{aligned}$$

Thus $\ker\left(L \otimes I_n\right) \cap \ker C = \{\mathbf{0}_{nN}\}$. Since $L \otimes I_n \succeq 0$ and $C \succeq 0$ and their kernels intersect trivially, $x^\top \mathcal{H} x = x^\top \left(\left(L \otimes I_n\right) + C\right) x > 0$ for all $x \neq \mathbf{0}_{nN}$. □

5.3.2 Closed-Loop Error System

Let $z \in \mathbb{R}^\varphi$ denote the concatenated state vector $z \triangleq \begin{bmatrix} e^\top & r^\top & \tilde{\eta}^\top & \tilde{r}^\top & \rho^\top & \tilde{\theta}^\top \end{bmatrix}^\top$, where $\varphi \triangleq (5n + p)\, N$. The evolution of z is governed by the initial value problem

$$\dot{z} = \mathtt{f}\left(z, t\right), \; z\left(t_0\right) = z_0, \tag{5-31}$$

where $z_0 \in \mathbb{R}^{\varphi}$ is the initial state. Using (5-7), (5-22), (5-27), (5-28), and (5-30), the vector field $\mathtt{f} : \mathbb{R}^{\varphi} \times \mathbb{R}_{\geq 0} \to \mathbb{R}^{\varphi}$ is defined as

$$\mathtt{f}(z,t) = \begin{bmatrix} r - k_1 e \\ \left(\begin{array}{c} \frac{\partial \widehat{\Psi}}{\partial \hat{\theta}} \tilde{\theta} - k_2 \mathcal{H} r + k_2 \tilde{r} + k_1 r \\ + \left(\left(k_1^2 + 1 \right) C - k_1^2 I_{nN} \right) e \end{array} \right) \\ -\omega + \tilde{f} + R + \varepsilon \\ \tilde{r} - k_3 \tilde{\eta} - \rho \\ \left(\begin{array}{c} \mathcal{H} \left(\frac{\partial \widehat{\Psi}}{\partial \hat{\theta}} \tilde{\theta} + R + \varepsilon - \omega + \tilde{f} \right) \\ + \left(k_3 + k_4 \right) \rho - \tilde{\eta} - k_4 \tilde{r} \end{array} \right) \\ \tilde{\eta} - \left(k_3 + k_4 \right) \tilde{r} - k_5 \rho \\ - \begin{bmatrix} \operatorname{proj}(\Theta_1)^{\top} & \cdots & \operatorname{proj}(\Theta_N)^{\top} \end{bmatrix}^{\top} \end{bmatrix}. \tag{5-32}$$

Since the projection operator in (5-29) is locally Lipschitz in its arguments [52, Lemma E.1] and the remaining terms in (5-32) are locally Lipschitz in z and continuous in t, $\mathtt{f}$ is locally Lipschitz in z and continuous in t. Therefore, by the Picard-Lindelöf theorem (e.g., [7, Chapter 1, Theorem 3.1]), for every $z_0 \in \mathbb{R}^{\varphi}$ there exists a unique maximal solution.

Since the universal approximation property of the ResNet holds only on the compact domain Ω, the subsequent stability analysis requires ensuring $\kappa_i(t) \in \Omega$ for all $i \in \mathcal{V}$ and for all $t \geq t_0$. This is achieved by yielding a stability result which constrains z to a compact domain. Consider the Lyapunov function candidate $V : \mathbb{R}^{\varphi} \to \mathbb{R}_{\geq 0}$ defined as

$$V(z) \triangleq \frac{1}{2} z^{\top} P z, \tag{5-33}$$

where $P \triangleq \operatorname{blkdiag}\left\{ I_{5nN}, \Gamma^{-1} \right\} \in \mathbb{R}^{\varphi \times \varphi}$ and $\Gamma \triangleq \operatorname{blkdiag}\{\Gamma_1, \ldots, \Gamma_N\} \in \mathbb{R}^{pN \times pN}$. By the Rayleigh quotient, (5-33) satisfies

$$\frac{1}{2} \lambda_1 \|z\|^2 \leq V(z) \leq \frac{1}{2} \lambda_{\varphi} \|z\|^2, \tag{5-34}$$

where $\lambda_1 \triangleq \lambda_{\min}\{P\} = \min\left\{1, \lambda_{\min}\left\{\Gamma^{-1}\right\}\right\}$ and $\lambda_\varphi \triangleq \lambda_{\max}\{P\} = \max\left\{1, \lambda_{\max}\left\{\Gamma^{-1}\right\}\right\}$. Furthermore, for any $x \neq \mathbf{0}_{nN}$, since $x^\top \mathcal{H} x = x^\top \mathcal{H}_{\text{Sym}} x$ with $\mathcal{H}_{\text{Sym}} \triangleq \frac{1}{2}\left(\mathcal{H} + \mathcal{H}^\top\right)$, the Rayleigh quotient yields

$$\lambda_{\min}\left\{\mathcal{H}_{\text{Sym}}\right\}\|x\|^2 \leq x^\top \mathcal{H} x \leq \lambda_{\max}\left\{\mathcal{H}_{\text{Sym}}\right\}\|x\|^2. \tag{5-35}$$

By the strict positivity of $x^\top \mathcal{H} x$ established by Theorem 5.1, it follows that $\mathcal{H}_{\text{Sym}} \succ 0$, hence $0 < \lambda_{\min}\left\{\mathcal{H}_{\text{Sym}}\right\} \leq \lambda_{\max}\left\{\mathcal{H}_{\text{Sym}}\right\}$. Similar to (5-17), define the regularized interaction matrix $\mathcal{J} \triangleq (L + \Xi) \otimes I_p \in \mathbb{R}^{pN \times pN}$ where $\Xi \triangleq \text{diag}\{\xi_1, \ldots, \xi_N\} \in \mathbb{R}^{N \times N}$.

Assumption 5.8. *For each $i \in \mathcal{V}$, ξ_i is selected to satisfy $\xi_i + \frac{1}{2}\left(d_i^{in} - d_i^{out}\right) > 0$.*

For any $x \neq \mathbf{0}_{pN}$, since $x^\top \mathcal{J} x = x^\top \mathcal{J}_{\text{Sym}} x$ with $\mathcal{J}_{\text{Sym}} \triangleq \frac{1}{2}\left(\mathcal{J} + \mathcal{J}^\top\right)$, the Rayleigh quotient yields

$$\lambda_{\min}\left\{\mathcal{J}_{\text{Sym}}\right\}\|x\|^2 \leq x^\top \mathcal{J} x \leq \lambda_{\max}\left\{\mathcal{J}_{\text{Sym}}\right\}\|x\|^2. \tag{5-36}$$

By the same estimate used in the first part of the proof of Theorem 5.1 and under Assumption 5.8, $x^\top \mathcal{J} x \geq \sum_{i=1}^N \left(\xi_i + \frac{1}{2}\left(d_i^{\text{in}} - d_i^{\text{out}}\right)\right)\|x_i\|^2 > 0$, so $\mathcal{J}_{\text{Sym}} \succ 0$ and every eigenvalue λ of $\mathcal{J}$ satisfies $\text{Re}(\lambda) > 0$. Consequently, $0 < \lambda_{\min}\left\{\mathcal{J}_{\text{Sym}}\right\} \leq \lambda_{\max}\left\{\mathcal{J}_{\text{Sym}}\right\}$.

By the mean value theorem (e.g., [37, Theorem 3.9]) and Taylor's theorem (e.g., [37, Theorem 4.7]), there exists a polynomial $\mathtt{p}_0(\|\kappa\|) = a_2\|\kappa\|^2 + a_1\|\kappa\| + a_0$ with $a_2, a_1, a_0 \in \mathbb{R}_{>0}$ such that $\left\|R\left(\kappa, \tilde{\theta}\right)\right\| \leq \mathtt{p}_0(\|\kappa\|)\left\|\tilde{\theta}\right\|^2$. Similarly, there exists a polynomial $\mathtt{p}_1(\|\kappa\|) = a_5\|\kappa\| + a_4$ with $a_5, a_4 \in \mathbb{R}_{>0}$ such that $\left\|\frac{\partial \widehat{\Psi}}{\partial \hat{\theta}}\right\| \leq \mathtt{p}_1(\|\kappa\|)\left\|\tilde{\theta}\right\|^2$.

By Assumption 5.2, there exist constants $\mathtt{L}_j \in \mathbb{R}_{\geq 0}$ such that $\left\|f_j(\mathtt{x}_1, \mathtt{y}_1) - f_j(\mathtt{x}_2, \mathtt{y}_2)\right\| \leq \mathtt{L}_j\sqrt{\|\mathtt{x}_1 - \mathtt{x}_2\|^2 + \|\mathtt{y}_1 - \mathtt{y}_2\|^2}$ for all $\mathtt{x}_1, \mathtt{x}_2, \mathtt{y}_1, \mathtt{y}_2 \in \mathbb{R}^n$ and $j \in \overline{\mathcal{V}}$. Based on the subsequent analysis, let $\mathtt{L} \triangleq \max_{j \in \overline{\mathcal{V}}} \mathtt{L}_j$, $\overline{\omega} \triangleq \max_{i \in \mathcal{V}}\{\overline{\omega}_i\}$, and define $\delta \triangleq N\left(\frac{5}{4k_2\lambda_{\min}\left\{\mathcal{H}_{\text{Sym}}\right\}} + \frac{\sigma_{\max}^2\{\mathcal{H}\}}{k_4}\right)\left(\overline{\varepsilon} + \overline{\omega} + \mathtt{L}\left(\left(\sqrt{2} + 2\right)\overline{q}_0 + 4\,\overline{\dot{q}}_0\right)\right)^2 + \frac{k_6 N \overline{\theta}^2 \sigma_{\max}^2\{\mathcal{J}\}}{2\lambda_{\min}\left\{\mathcal{J}_{\text{Sym}}\right\}}$ and $k_{\min} \triangleq \min\left\{\frac{1}{4}k_1, \frac{1}{5}k_2\lambda_{\min}\left\{\mathcal{H}_{\text{Sym}}\right\} - \frac{1}{k_1}\left(\left(k_1^2 + 1\right)\left(\lambda_{\max}\{C\} + 1\right) + 2\mathtt{L}\left(1 + k_1\right)\left(1 + \sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right)\right)^2 - k_1 - \mathtt{L}\left(\left(2 + \sigma_{\max}\{\mathcal{H}\}\right)\left(1 + \sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right) + 2\sqrt{2}\right), \frac{1}{4}k_3, \frac{1}{4}k_4 - 4\mathtt{L}^2\sigma_{\max}^2\{\mathcal{H}\}\left(\frac{1}{k_1}\left(1 + k_1\right)^2\left(1 + \sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right)^2 - 2k_3\right) - \mathtt{L}\sigma_{\max}\{\mathcal{H}\}\left(\left(3 + \sigma_{\max}\{\mathcal{H}\}\right)\sqrt{2} + 1\right) - \frac{5k_2}{4\lambda_{\min}\left\{\mathcal{H}_{\text{Sym}}\right\}}, \frac{1}{2}k_5 - \mathtt{L}\sqrt{2}\left(1 + \sigma_{\max}\{\mathcal{H}\}\right), \frac{1}{4}k_6\,\lambda_{\min}\left\{\mathcal{J}_{\text{Sym}}\right\}\right\}$, where

$\mathrm{p}\left(\|\kappa\|, [k_i]_{i\in[5]}\right) \triangleq N\overline{\theta}^2 \left(\frac{5}{k_2\lambda_{\min}\{\mathcal{H}_{\mathrm{Sym}}\}} + 4\sigma_{\max}^2\{\mathcal{H}\}\left(\frac{1}{k_3} + \frac{1}{k_4}\right)\right)\mathrm{p}_0^2(\|\kappa\|) +$
$4N^2\overline{\theta}^4 (4\,k_1 (\sigma_{\max}^2\{\mathcal{H}\} + \sigma_{\max}\{\mathcal{H}\}\sigma_{\max}\{A\} + \lambda_{\max}\{C\} + 2)^2 + \frac{5}{k_2\lambda_{\min}\{\mathcal{H}_{\mathrm{Sym}}\}}(\sigma_{\max}^2\{\mathcal{H}\} +$
$\sigma_{\max}\{\mathcal{H}\}\sigma_{\max}\{A\} + 1)^2 + 4\sigma_{\max}^2\{A\}(4\,k_3 + \frac{1}{k_4} + \frac{2}{k_5})\Big)\mathrm{p}_1^2(\|\kappa\|)$.

Since $\mathrm{p}_0(\cdot)$ and $\mathrm{p}_1(\cdot)$ are strictly increasing, $\mathrm{p}(\cdot)$ is strictly increasing and there exists a strictly increasing $\varrho : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\varrho(\|z\|) \triangleq \frac{\mathrm{p}^2(\|\kappa\|,[k_i]_{i\in[5]})}{\lambda_{\min}\{\mathcal{J}_{\mathrm{Sym}}\}}$. Define $\overline{\varrho}(\cdot) \triangleq \varrho(\cdot) - \varrho(0)$, where $\overline{\varrho}$ is strictly increasing and invertible. The region to which the state trajectory is constrained is defined as

$$\mathcal{D} \triangleq \left\{\iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq \overline{\varrho}^{-1}\left(k_6\left(k_{\min} - \lambda_V\right) - \varrho(0)\right)\right\}, \tag{5-37}$$

where $\lambda_V \in \mathbb{R}_{>0}$ is a user-defined parameter which controls the desired rate of convergence. The compact domain over which the universal approximation property must hold is selected as

$$\Omega \triangleq \left\{\iota \in \mathbb{R}^{2n} : \|\iota\| \leq ((k_1+2)\lambda_{\max}\{\mathcal{H}\} + k_3 + 2)\cdot\overline{\varrho}^{-1}\left(k_6\left(k_{\min} - \lambda_V\right) - \varrho(0)\right)\right\}. \tag{5-38}$$

For the dynamical system described by (5-31), the set of initial conditions is defined as

$$\mathcal{S} \triangleq \left\{\iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq +\sqrt{\frac{\lambda_1}{\lambda_\varphi}}\,\overline{\varrho}^{-1}\left(k_6\left(k_{\min} - \lambda_V\right) - \varrho(0)\right)\right\} - \sqrt{\frac{\delta}{\lambda_V}}\Bigg\}, \tag{5-39}$$

and the uniform ultimate bound is defined as

$$\mathcal{U} \triangleq \left\{\iota \in \mathbb{R}^{\varphi} : \|\iota\| \leq \sqrt{\frac{\lambda_\varphi\delta}{\lambda_1\lambda_V}}\right\}. \tag{5-40}$$

Theorem 5.2. *Consider the dynamical system described by* (5-1) *and* (5-2)*. For any $z_0 \in \mathcal{S}$, the observer* (5-14)*, controller* (5-15)*, and adaptation law* (5-16) *ensure that z uniformly ultimately bounded and exponentially converges to $\mathcal{U}$ with estimate*

$$\|z(t)\| \leq \sqrt{\frac{\lambda_\varphi}{\lambda_1}}\sqrt{\|z_0\|^2\,\mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} + \frac{\delta}{\lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)}\right)},$$

for all $t \geq t_0$*, provided that the sufficient gain condition* $k_{\min} > \lambda_V + \frac{1}{k_6}\varrho\left(\sqrt{\frac{\delta}{\lambda_V}}\left(\sqrt{\frac{\lambda_\varphi}{\lambda_1}} + \frac{\lambda_\varphi}{\lambda_1}\right)\right)$ *is satisfied and Assumptions 5.1-5.8 hold. Moreover,* $\mathcal{U} \subset \mathcal{S} \subset \mathcal{D}$*,* $\kappa_i(t) \in \Omega$ *for all* $t \geq t_0$ *and* $i \in \mathcal{V}$*, and the solution* z *is complete.*

Proof. Taking the total derivative of (5-33) along the trajectories of (5-31) yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V(z(t)) = {} & e^\top(t)\,\dot{e}(t) + r^\top(t)\,\dot{r}(t) + \tilde{\eta}^\top(t)\,\dot{\tilde{\eta}}(t) + \tilde{r}^\top(t)\,\dot{\tilde{r}}(t) \\ & + \rho^\top(t)\,\dot{\rho}(t) + \tilde{\theta}^\top(t)\,\Gamma^{-1}\dot{\tilde{\theta}}(t).\end{aligned} \tag{5-41}$$

Substituting (5-32) into (5-41) and using (5-25) yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V(z(t)) = {} & -k_1 e^\top(t)\,e(t) - k_2 r^\top(t)\,\mathcal{H}r(t) - k_3\tilde{\eta}^\top(t)\,\tilde{\eta}(t) - k_4\tilde{r}^\top(t)\,\tilde{r}(t) - k_5\rho^\top(t)\,\rho(t) \\ & + k_2 r^\top(t)\,\tilde{r}(t) + k_1 r^\top(t)\,r(t) + r^\top(t)\left(\left(k_1^2+1\right)C - \left(k_1^2-1\right)I_{nN}\right)e(t) \\ & + \left(r^\top(t) + \tilde{r}^\top(t)\,\mathcal{H}\right)\left(R\left(\kappa(t),\tilde{\theta}(t)\right) + \varepsilon(t) - \omega(t) + \tilde{f}(t)\right) \\ & - \sum_{i\in\mathcal{V}}\tilde{\theta}_i^\top(t)\,\Gamma_i^{-1}(t)\,\mathrm{proj}\left(\Theta_i(t)\right) + \tilde{\theta}^\top(t)\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\right)^\top\left(r(t) + \mathcal{H}^\top\tilde{r}(t)\right).\end{aligned} \tag{5-42}$$

Using (5-21), the ensemble representations of (5-7)-(5-8) and (5-12), the fact that $\mathcal{H} = (D \otimes I_n) - (A \otimes I_n) + C$ and the symmetric properties of D and C yields

$$\begin{aligned}\tilde{\theta}^\top\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\right)^\top\left(r + \mathcal{H}^\top\tilde{r}\right) = {} & \tilde{\theta}^\top\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\right)^\top\left(k_1 e + \left((D\otimes I_n) + C\right)\left(k_3\tilde{\eta} + \rho - \hat{\zeta}\right)\right) \\ & + \tilde{\theta}^\top\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\right)^\top\left(\dot{e} + \mathcal{H}^\top\zeta - \left(A^\top\otimes I_n\right)\left(k_3\tilde{\eta} + \rho - \hat{\zeta}\right)\right).\end{aligned} \tag{5-43}$$

Invoking [52, Appendix E, Lemma E.1.IV] and using (5-4) and the ensemble representation of (5-25) yields

$$\begin{aligned}-\sum_{i\in\mathcal{V}}\tilde{\theta}_i^\top\Gamma_i^{-1}\mathrm{proj}\left(\Theta_i\right) \leq {} & -\tilde{\theta}^\top\left(\frac{\partial\widehat{\Psi}}{\partial\hat{\theta}}\right)^\top\left(k_1 Ce + \left((D\otimes I_n) + C\right)\left(k_3\tilde{\eta} + \rho - \hat{\zeta}\right)\right) \\ & + k_6\tilde{\theta}^\top\mathcal{J}\theta^* - k_6\tilde{\theta}^\top\mathcal{J}\tilde{\theta},\end{aligned} \tag{5-44}$$

where $\theta^* \triangleq \left[\theta_i^*\right]_{i\in\mathcal{V}}$. By construction, $\theta_i^* \in \mathfrak{V}_i$, for all $i \in \mathcal{V}$. With $\overline{\theta} = \max_{i\in\mathcal{V}} \sup_{\vartheta\in\mathfrak{V}_i} \|\vartheta\|$, it follows that $\left\|\theta_i^*\right\| \leq \overline{\theta}$ for all $i \in \mathcal{V}$, hence $\|\theta^*\| \leq \overline{\theta}\sqrt{N}$. Likewise, $\|\varepsilon\| \leq \overline{\varepsilon}\sqrt{N}$, and by Assumption 5.4, $\|\omega\| \leq \overline{\omega}\sqrt{N}$. Invoking (5-35), (5-36), (5-43) and (5-44), together with the triangle and Cauchy-Schwarz inequalities yields that (5-42) is bounded above as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}V\left(z\left(t\right)\right) \leq & -k_1 \left\|e\left(t\right)\right\|^2 - k_2\lambda_{\min}\left\{\mathcal{H}_{\mathrm{Sym}}\right\}\left\|r\left(t\right)\right\|^2 - k_3\left\|\tilde{\eta}\left(t\right)\right\|^2 - k_4\left\|\tilde{r}\left(t\right)\right\|^2 - k_5\left\|\rho\left(t\right)\right\|^2 \\
& - k_6\lambda_{\min}\left\{\mathcal{J}_{\mathrm{Sym}}\right\}\left\|\tilde{\theta}\left(t\right)\right\|^2 + k_2\left\|r\left(t\right)\right\|\left\|\tilde{r}\left(t\right)\right\| + k_1\left\|r\left(t\right)\right\|^2 \\
& + \left(k_1^2+1\right)\left(\lambda_{\max}\left\{C\right\}+1\right)\left\|e\left(t\right)\right\|\left\|r\left(t\right)\right\| + \sqrt{N}k_6\overline{\theta}\sigma_{\max}\left\{\mathcal{J}\right\}\left\|\tilde{\theta}\left(t\right)\right\| \\
& + \left(\left\|r\left(t\right)\right\| + \sigma_{\max}\left\{\mathcal{H}\right\}\left\|\tilde{r}\left(t\right)\right\|\right)\left(\mathrm{p}_0\left(\left\|\kappa\left(t\right)\right\|\right)\left\|\tilde{\theta}\left(t\right)\right\|^2 + \sqrt{N}\overline{\varepsilon} + \sqrt{N}\overline{\omega} + \left\|\tilde{f}\left(t\right)\right\|\right) \\
& + k_1\mathrm{p}_1\left(\left\|\kappa\left(t\right)\right\|\right)\left(1+\lambda_{\max}\left\{C\right\}\right)\left\|\tilde{\theta}\left(t\right)\right\|^3\left\|e\left(t\right)\right\| \\
& + \mathrm{p}_1\left(\left\|\kappa\left(t\right)\right\|\right)\left\|\tilde{\theta}\left(t\right)\right\|^3\left(\sigma_{\max}\left\{A\right\}\left(k_3\left\|\tilde{\eta}\left(t\right)\right\| + \left\|\rho\left(t\right)\right\| + \left\|\hat{\zeta}\left(t\right)\right\|\right)\right) \\
& + \mathrm{p}_1\left(\left\|\kappa\left(t\right)\right\|\right)\left\|\tilde{\theta}\left(t\right)\right\|^3\left(\left\|\dot{e}\left(t\right)\right\| + \sigma_{\max}\left\{\mathcal{H}\right\}\left\|\zeta\left(t\right)\right\|\right)
\end{aligned} \tag{5-45}$$

Using (5-18) yields

$$\left\|\eta\right\| \leq \sigma_{\max}\left\{\mathcal{H}\right\}\left\|e\right\|. \tag{5-46}$$

Using (5-19) yields

$$\left\|\zeta\right\| \leq \sigma_{\max}\left\{\mathcal{H}\right\}\left(\left\|r\right\| + k_1\left\|e\right\|\right). \tag{5-47}$$

Using the ensemble representation of (5-7), (5-8), (5-18), (5-19), and (5-47) yields

$$\left\|\hat{\zeta}\right\| \leq k_1\sigma_{\max}\left\{\mathcal{H}\right\}\left\|e\right\| + \sigma_{\max}\left\{\mathcal{H}\right\}\left\|r\right\| + \left\|\tilde{r}\right\| + k_3\left\|\tilde{\eta}\right\| + \left\|\rho\right\|. \tag{5-48}$$

Using the ensemble representation of (5-7) yields

$$\left\|\dot{e}\right\| \leq k_1\left\|e\right\| + \left\|r\right\|. \tag{5-49}$$

Using Assumption 5.2 together with the definition of $\tilde{f}_i$, the triangle inequality, Assumption 5.1, and (5-5), (5-46) and (5-48), it follows that

$$\begin{aligned}\left\|\tilde{f}\right\| \leq 2\mathrm{L}\left(1+k_1+(1+k_1)\sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right)\|e\|+2\mathrm{L}\left(\sqrt{2}\sigma_{\max}\{\mathcal{H}\}+1\right)\|r\| \\ +2\mathrm{L}\sqrt{2}\|\tilde{r}\|+2k_3\mathrm{L}\sqrt{2}\|\tilde{\eta}\|+2\mathrm{L}\sqrt{2}\|\rho\|+\mathrm{L}\sqrt{N}\left(\sqrt{2}+2\right)\overline{q}_0+4\mathrm{L}\sqrt{N}\bar{q}_0.\end{aligned} \tag{5-50}$$

Thus using (5-46)-(5-50) yields that (5-45) is bounded above as

$$\frac{\mathrm{d}}{\mathrm{d}t}V\left(z\left(t\right)\right) \leq V_e\left(t\right)+V_r\left(t\right)+V_{\tilde{\eta}}\left(t\right)+V_{\tilde{r}}\left(t\right)+V_{\rho}\left(t\right)+V_{\tilde{\theta}}\left(t\right)+V_y\left(t\right), \tag{5-51}$$

where

$$
\begin{aligned}
V_e(t) &\triangleq -k_1 \|e(t)\|^2 + \sigma_{\max}\{\mathcal{H}\} 2\mathrm{L}\left(1 + k_1 + (1+k_1)\sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right)\|e(t)\|\,\|\tilde{r}(t)\| \\
&\quad + \left(\left(k_1^2+1\right)(\lambda_{\max}\{C\}+1) + 2\mathrm{L}\left(1 + k_1 + (1+k_1)\sqrt{2}\sigma_{\max}\{\mathcal{H}\}\right)\right)\|e(t)\|\,\|r(t)\| \\
&\quad + k_1\left(\left(\sigma_{\max}^2\{\mathcal{H}\} + \lambda_{\max}\{C\} + 2\right) + \sigma_{\max}\{\mathcal{H}\}\sigma_{\max}\{A\}\right)\mathrm{p}_1(\|\kappa\|)\left\|\tilde{\theta}(t)\right\|^3\|e(t)\|, \\
V_r(t) &\triangleq -k_2\lambda_{\min}\left\{\mathcal{H}_{\mathrm{Sym}}\right\}\|r(t)\|^2 + k_2\|r(t)\|\,\|\tilde{r}(t)\| + \left(k_1 + 2\mathrm{L}\left(\sqrt{2}\sigma_{\max}\{\mathcal{H}\}+1\right)\right)\|r(t)\|^2 \\
&\quad + \mathrm{p}_0(\|\kappa(t)\|)\|r(t)\|\left\|\tilde{\theta}(t)\right\|^2 + \sqrt{N}\left(\overline{\varepsilon} + \overline{\omega} + \mathrm{L}\left(\left(\sqrt{2}+2\right)\overline{q}_0 + 4\overline{\dot{q}}_0\right)\right)\|r(t)\| \\
&\quad + \left(\sigma_{\max}^2\{\mathcal{H}\} + \sigma_{\max}\{\mathcal{H}\}\sigma_{\max}\{A\} + 1\right)\mathrm{p}_1(\|\kappa(t)\|)\left\|\tilde{\theta}(t)\right\|^3\|r(t)\|, \\
V_{\tilde{\eta}}(t) &\triangleq -k_3\|\tilde{\eta}(t)\|^2 + 2k_3\sigma_{\max}\{A\}\mathrm{p}_1(\|\kappa(t)\|)\left\|\tilde{\theta}(t)\right\|^3\|\tilde{\eta}(t)\| + 2k_3\mathrm{L}\sqrt{2}\|\tilde{\eta}(t)\|\,\|r(t)\| \\
&\quad + 2k_3\mathrm{L}\sqrt{2}\sigma_{\max}\{\mathcal{H}\}\|\tilde{\eta}(t)\|\,\|\tilde{r}(t)\|, \\
V_{\tilde{r}}(t) &\triangleq -k_4\|\tilde{r}(t)\|^2 + 2\mathrm{L}\sigma_{\max}\{\mathcal{H}\}\sqrt{2}\|\tilde{r}(t)\|^2 \\
&\quad + \sqrt{N}\sigma_{\max}\{\mathcal{H}\}\left(\overline{\varepsilon} + \overline{\omega} + \mathrm{L}\left(\left(\sqrt{2}+2\right)\overline{q}_0 + 4\overline{\dot{q}}_0\right)\right)\|\tilde{r}(t)\| \\
&\quad + \left(\sigma_{\max}\{A\}\mathrm{p}_1(\|\kappa(t)\|)\left\|\tilde{\theta}(t)\right\|^3 + \sigma_{\max}\{\mathcal{H}\}\mathrm{p}_0(\|\kappa(t)\|)\left\|\tilde{\theta}\right\|^2\right)\|\tilde{r}(t)\|, \\
V_{\rho}(t) &\triangleq -k_5\|\rho(t)\|^2 + 2\sigma_{\max}\{A\}\mathrm{p}_1(\|\kappa(t)\|)\left\|\tilde{\theta}(t)\right\|^3\|\rho(t)\|, \\
V_{\tilde{\theta}}(t) &\triangleq -k_6\lambda_{\min}\left\{\mathcal{J}_{\mathrm{Sym}}\right\}\left\|\tilde{\theta}(t)\right\|^2 + \sqrt{N}k_6\overline{\theta}\sigma_{\max}\{\mathcal{J}\}\left\|\tilde{\theta}(t)\right\|, \\
V_y(t) &\triangleq 2\mathrm{L}\left(\sqrt{2} + \sigma_{\max}\{\mathcal{H}\}\left(\sqrt{2}\sigma_{\max}\{\mathcal{H}\}+1\right)\right)\|r(t)\|\,\|\tilde{r}(t)\| + 2\mathrm{L}\sqrt{2}\|\rho(t)\|\,\|r(t)\| \\
&\quad + 2\mathrm{L}\sigma_{\max}\{\mathcal{H}\}\sqrt{2}\|\rho(t)\|\,\|\tilde{r}(t)\|.
\end{aligned}
$$

Applying Young's inequality, completing the square, and using the bounds $\left\|\tilde{\theta}\right\| \leq 2\sqrt{N\theta}$ and $\left\|\tilde{\theta}\right\| \leq \|z\|$, together with the definitions of $\varrho(\|z\|)$, $k_{\min}$, and δ, implies that (5-51) is bounded above by

$$
\frac{\mathrm{d}}{\mathrm{d}t}V(z(t)) \leq -\left(k_{\min} - \frac{\varrho(\|z\|)}{k_6}\right)\|z\|^2 + \delta. \tag{5-52}
$$

It is now shown, by contradiction, that the maximal solution to (5-31) exists for all time and remains in the interior of $\mathcal{D}$. Fix any $z_0 \in \mathcal{S}$ and let $z : [t_0, T_{\max}) \to \mathbb{R}^{\varphi}$ be the corresponding

unique maximal solution of (5-31). Let $\mathcal{I} \triangleq \{t \in [t_0, T_{\max}) : z(\tau) \in \mathcal{D} \text{ for all } \tau \in [t_0, t]\}$, where $T_{\max} \in (t_0, \infty]$ denotes the supremum of all times $T > t_0$ for which a solution exists on $[t_0, T)$. Since $z_0 \in \mathcal{S} \subset \mathcal{D}$ and $t \mapsto z(t)$ is continuous, the set $\mathcal{I}$ is non-empty and contains an interval $[t_0, t_0 + \epsilon)$ for some $\epsilon > 0$.

For any $t \in \mathcal{I}$, the construction of $\mathcal{D}$ gives $\overline{\varrho}(\|z(t)\|) \leq k_6(k_{\min} - \lambda_V) - \varrho(0)$. Hence, using $\overline{\varrho}(\cdot) = \varrho(\cdot) - \varrho(0)$ implies $k_{\min} - \frac{\varrho(\|z\|)}{k_6} \geq \lambda_V$. Therefore, combining (5-34) with (5-52) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} V(z(t)) \leq -\frac{2\lambda_V}{\lambda_\varphi} V(z(t)) + \delta, \tag{5-53}$$

for all $t \in \mathcal{I}$. Solving the differential inequality given by (5-53) over $\mathcal{I}$ yields

$$V(z(t)) \leq V(z(t_0))\, \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} + \frac{\lambda_\varphi \delta}{2\lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)}\right) \tag{5-54}$$

for all $t \in \mathcal{I}$. Applying (5-34) to (5-54) yields

$$\|z(t)\| \leq \sqrt{\frac{\lambda_\varphi}{\lambda_1}\|z(t_0)\|^2\, \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} + \frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}\left(1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)}\right)}, \tag{5-55}$$

for all $t \in \mathcal{I}$. Using (5-39), since $z_0 \in \mathcal{S}$, $\|z_0\| \leq \sqrt{\frac{\lambda_1}{\lambda_\varphi}}\overline{\varrho}^{-1}(k_6(k_{\min} - \lambda_V) - \rho(0)) - \sqrt{\frac{\delta}{\lambda_V}}$. Substituting this expression into (5-55) and using the facts that $\mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} \leq 1$ and $1 - \mathrm{e}^{-\frac{2\lambda_V}{\lambda_\varphi}(t-t_0)} < 1$ for all $t \geq t_0$ yield the strict bound $\|z(t)\| < \overline{\varrho}^{-1}(k_6(k_{\min} - \lambda_V) - \rho(0))$ for all $t \in \mathcal{I}$. Thus, by (5-37), $z(t) \in \operatorname{int}(\mathcal{D})$ for all $t \in \mathcal{I}$. Now, assume for contradiction that the maximal time of existence is finite, i.e., $\sup(\mathcal{I}) < T_{\max}$. If $\sup(\mathcal{I}) < T_{\max}$, continuity of $z(\cdot)$ would imply $z(\sup(\mathcal{I})) \in \partial\mathcal{D}$, contradicting $z(t) \in \operatorname{int}(\mathcal{D})$ for all $t \in \mathcal{I}$. Hence, $\sup(\mathcal{I}) = T_{\max}$ and $\mathcal{I} = [t_0, T_{\max})$, i.e., $z(t) \in \mathcal{D}$ for all $t \in [t_0, T_{\max})$.

Since $z(t) \in \mathcal{D}$ for all $t \in [t_0, T_{\max})$ and $\mathcal{D}$ is compact, $\sup_{t \in [t_0, T_{\max})} \|z(t)\| < \infty$. By [7, Chapter 1, Theorem 4.1], the solution extends from $[t_0, T_{\max})$ to $[t_0, \infty)$. Therefore, the solution

exists for all $t \geq t_0$ with $z(t) \in \mathcal{D}$ for all $t \geq t_0$. Consequently, for all $z_0 \in \mathcal{S}$, (5-55) holds for every $t \geq t_0$.

As $t \to \infty$, the bound converges to $\|z(t)\| \leq \sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}$, i.e., the solution is ultimately bounded with ultimate bound $\mathcal{U}$ by (5-40). Next, recall that $k_{\min} > \lambda_V + \frac{1}{k_6} \varrho\left(\sqrt{\frac{\delta}{\lambda_V}}\left(\sqrt{\frac{\lambda_\varphi}{\lambda_1}} + \frac{\lambda_\varphi}{\lambda_1}\right)\right)$. Using the definition of $\overline{\varrho}$ and the fact that $\overline{\varrho}$ is invertible yields $\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}} < \sqrt{\frac{\lambda_1}{\lambda_\varphi}} \overline{\varrho}^{-1}\left(k_6 (k_{\min} - \lambda_V) - \varrho(0)\right) - \sqrt{\frac{\delta}{\lambda_V}}$, i.e., $\mathcal{U} \subset \mathcal{S}$. Furthermore, since $\frac{\lambda_1}{\lambda_\varphi} \leq 1$ and $\sqrt{\frac{\delta}{\lambda_V}} > 0$, strict monotonicity of ϱ gives $k_{\min} > \lambda_V + \frac{1}{k_6} \varrho\left(\sqrt{\frac{\delta}{\lambda_V}}\left(\sqrt{\frac{\lambda_\varphi}{\lambda_1}} + \frac{\lambda_\varphi}{\lambda_1}\right)\right) > \lambda_V + \frac{1}{k_6} \varrho\left(\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}\right)$ which implies $k_6 (k_{\min} - \lambda_V) - \varrho(0) > \overline{\varrho}\left(\sqrt{\frac{\lambda_\varphi \delta}{\lambda_1 \lambda_V}}\right) > 0$. Hence, $\sqrt{\frac{\lambda_1}{\lambda_\varphi}} \overline{\varrho}^{-1}\left(k_6 (k_{\min} - \lambda_V) - \varrho(0)\right) - \sqrt{\frac{\delta}{\lambda_V}} > 0$, so $\mathcal{S}$ has strictly positive radius i.e., $\mathcal{S}$ is nonempty. Thus, $\mathcal{U} \subset \mathcal{S} \subset \mathcal{D}$.

Using the definitions of κ_i, z, and (5-46) and (5-48) yields

$$\|\kappa_i(t)\| \leq \left((k_1 + 2)\, \sigma_{\max}\{\mathcal{H}\} + k_3 + 2\right) \|z(t)\|, \tag{5-56}$$

for all $t \geq t_0$. Since $z(t) \in \mathcal{D}$ for all $t \geq t_0$, the images of $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{D}$ under $z \mapsto \kappa_i$ lie in Ω by (5-38). Consequently, $\kappa_i(t) \in \Omega$ for all $t \geq t_0$, and the universal approximation property applied in (5-11) holds for all time.

Since (5-55) holds with constants λ_1, λ_φ, λ_V, δ chosen independently of t_0 and z_0, the right-hand side depends on time only through $(t - t_0)$. Hence, the exponential convergence to $\mathcal{U}$ is uniform (e.g., [76, Definition 1]).

Additionally, since $z(t) \in \mathcal{D}$ for all $t \geq t_0$, $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$. Since $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$, $e, r, \tilde{\eta}, \tilde{r}, \rho \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{nN}\right)$ and $\tilde{\theta} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^p\right)$. Thus, $\hat{\zeta} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{nN}\right)$ by (5-48). Similarly, since $q_0, \dot{q}_0 \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^n\right)$ by Assumption 5.1, using (5-5) and (5-7) yields that $q_i, \dot{q}_i \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^n\right)$ for all $i \in \mathcal{V}$. Hence, using Assumption 5.5 yields that $g^+ \in \mathcal{L}_\infty\left(\mathbb{R}^{nN} \times \mathbb{R}^{nN} \times \mathbb{R}_{\geq t_0}; \mathbb{R}^{\sum_{i \in \mathcal{V}} s_i \times nN}\right)$. Following (5-56) and the fact that $z \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^\varphi\right)$ yields that $\kappa_i \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{2n}\right)$ for all $i \in \mathcal{V}$. By the use of the projection operator, $\hat{\theta}_i \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{p_i}\right)$ for all $i \in \mathcal{V}$. Since $\left(\kappa_i, \hat{\theta}_i\right) \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{2n} \times \mathbb{R}^{p_i}\right)$, $\Psi_i \in \mathcal{L}_\infty\left(\mathbb{R}^{2n} \times \mathbb{R}^{p_i}; \mathbb{R}^n\right)$

for all $i \in \mathcal{V}$. Thus, by (5-20), $u \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{nN}\right)$. Since $u, \hat{\zeta}, \tilde{\eta}, \rho \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{nN}\right)$, $\dot{\tilde{\eta}}, \dot{\hat{\zeta}} \in \mathcal{L}_\infty\left(\mathbb{R}_{\geq t_0}; \mathbb{R}^{nN}\right)$ by (5-21). Thus, all implemented signals are bounded. □

5.4 Spacecraft Servicing Simulation

This section validates the empirical performance of the distributed observer, controller, and adaptive update law given by (5-14)-(5-16). The scenario considers a network of $N = 8$ servicer spacecraft tasked with tracking a single defunct spacecraft during an approach and servicing operation.

The defunct spacecraft follows an elliptical Keplerian orbit about Earth (central body), modeled as a point mass with standard gravitational parameter $\mu = 3.986 \times 10^{14}\ \mathrm{m}^3 \cdot \mathrm{s}^{-2}$. Relative motion is expressed in the target-centered radial-tangential-normal (RTN) frame, also known as the local-vertical-local-horizontal (LVLH) frame, with axes $\{R, T, N\}$.

For radar-based rendezvous, line-of-sight (LOS) spherical coordinates $(\sigma_i, \gamma_i, \phi_i)$ for servicer i are defined from the RTN Cartesian relative position $\mathbf{r}_i^{\mathrm{RTN}} \triangleq \begin{bmatrix} r_i^x & r_i^y & r_i^z \end{bmatrix}^\top \in \mathbb{R}^3$ (m) by $r_i^x = \sigma_i \cos\phi_i \cos\gamma_i$, $r_i^y = \sigma_i \cos\phi_i \sin\gamma_i$, and $r_i^z = \sigma_i \sin\phi_i$, where σ_i is range (m), γ_i is azimuth (rad) measured in the RT-plane from $+R$ toward $+T$, and ϕ_i is elevation (rad) measured from the RT-plane toward $+N$ (e.g., [10]). The inverse map is $\sigma_i = \sqrt{\left(r_i^x\right)^2 + \left(r_i^y\right)^2 + \left(r_i^z\right)^2}$, $\gamma_i = \operatorname{atan2}\left(r_i^y, r_i^x\right)$, and $\phi_i = \arcsin\left(r_i^z/\sigma_i\right)$.

Define the spherical state $q_i \triangleq \begin{bmatrix} \sigma_i & \gamma_i & \phi_i \end{bmatrix}^\top \in \mathbb{R}^3$ and LOS-resolved control $u_i \triangleq \begin{bmatrix} u_i^\sigma & u_i^\gamma & u_i^\phi \end{bmatrix}^\top \in \mathbb{R}^3\left(\mathrm{m} \cdot \mathrm{s}^{-2}\right)$. The spherical rates are $\dot{q}_i = \begin{bmatrix} \dot{\sigma}_i & \dot{\gamma}_i & \dot{\phi}_i \end{bmatrix}^\top \in \mathbb{R}^3$ with units $\left(\mathrm{m} \cdot \mathrm{s}^{-1}, \mathrm{rad} \cdot \mathrm{s}^{-1}, \mathrm{rad} \cdot \mathrm{s}^{-1}\right)$, and $\ddot{q}_i = \begin{bmatrix} \ddot{\sigma}_i & \ddot{\gamma}_i & \ddot{\phi}_i \end{bmatrix}^\top \in \mathbb{R}^3$ with units $\left(\mathrm{m} \cdot \mathrm{s}^{-2}, \mathrm{rad} \cdot \mathrm{s}^{-2}, \mathrm{rad} \cdot \mathrm{s}^{-2}\right)$. The LOS orthonormal basis vectors are $\mathbf{e}_{\sigma_i} = \begin{bmatrix} (\cos\phi_i \cos\gamma_i) & (\cos\phi_i \sin\gamma_i) & \sin\phi_i \end{bmatrix}^\top$, $\mathbf{e}_{\gamma_i} = \begin{bmatrix} (-\sin\gamma_i) & (\cos\gamma_i) & 0 \end{bmatrix}^\top$, and $\mathbf{e}_{\phi_i} = \begin{bmatrix} (-\sin\phi_i \cos\gamma_i) & (-\sin\phi_i \sin\gamma_i) & (\cos\phi_i) \end{bmatrix}^\top$.

Let $B(q_i) \in \mathrm{SO}(3)$ denote the rotation matrix with columns $\mathbf{e}_{\sigma_i}, \mathbf{e}_{\gamma_i}, \mathbf{e}_{\phi_i}$, i.e., $B(q_i) \triangleq \begin{bmatrix} \mathbf{e}_{\sigma_i} & \mathbf{e}_{\gamma_i} & \mathbf{e}_{\phi_i} \end{bmatrix}$. The RTN control accelerations are $u_i^{\mathrm{RTN}} \triangleq \begin{bmatrix} u_i^x & u_i^y & u_i^z \end{bmatrix}^\top \in \mathbb{R}^3\left(\mathrm{m} \cdot \mathrm{s}^{-2}\right)$. Orthonormality implies $u_i^{\mathrm{RTN}} = B(q_i)\, u_i$ and

$u_i = B^\top(q_i)\, u_i^{\mathrm{RTN}}$. Hence, $u_i^\sigma, u_i^\gamma, u_i^\phi$ have units of $\mathrm{m}\cdot\mathrm{s}^{-2}$. When mapped through $g(q_i)$ (defined below), the angular components contribute $\mathrm{rad}\cdot\mathrm{s}^{-2}$.

Additional accelerations include gravitational perturbations and atmospheric drag, grouped as $\mathbf{P}_i^{\mathrm{RTN}}(q_i, \mathtt{r}, \eta) = \begin{bmatrix} \mathbf{P}_i^x & \mathbf{P}_i^y & \mathbf{P}_i^z \end{bmatrix}^\top \in \mathbb{R}^3\left(\mathrm{m}\cdot\mathrm{s}^{-2}\right)$ and $\mathbf{D}_i^{\mathrm{RTN}}(\dot{q}_i, \mathtt{r}, \eta) = \begin{bmatrix} \mathbf{D}_i^x & \mathbf{D}_i^y & \mathbf{D}_i^z \end{bmatrix}^\top \in \mathbb{R}^3\left(\mathrm{m}\cdot\mathrm{s}^{-2}\right)$, modeling effects such as the oblateness coefficient $J_2 \in \mathbb{R}_{>0}$ (dimensionless), third-body gravity, and atmospheric drag. Here, $\mathtt{r} \in \mathbb{R}\,(\mathrm{m})$ is the defunct spacecraft's orbital radius. Exogenous variability is parameterized by a vector of phases $\eta \in \mathbb{T}^k$ (the k-torus of angles in radians), evolving with constant rate $\dot{\eta} = \Omega \in \mathbb{R}^k\left(\mathrm{rad}\cdot\mathrm{s}^{-1}\right)$.

The RTN dynamics of servicer i relative to the defunct spacecraft (e.g., [87]) are

$$\begin{aligned}
\ddot{r}_i^x &= 2\tau\dot{r}_i^y + \dot{\tau}r_i^y + \tau^2 r_i^x + \frac{\mu}{\mathtt{r}^2} + \mathbf{P}_i^x + \mathbf{D}_i^x + u_i^x - \frac{\mu\left(\mathtt{r} + r_i^x\right)}{\left(\left(\mathtt{r} + r_i^x\right)^2 + \left(r_i^y\right)^2 + \left(r_i^z\right)^2\right)^{\frac{3}{2}}},\\
\ddot{r}_i^y &= -2\tau\dot{r}_i^x - \dot{\tau}r_i^x + \tau^2 r_i^y + \mathbf{P}_i^y + \mathbf{D}_i^y + u_i^y - \frac{\mu r_i^y}{\left(\left(\mathtt{r} + r_i^x\right)^2 + \left(r_i^y\right)^2 + \left(r_i^z\right)^2\right)^{\frac{3}{2}}},\\
\ddot{r}_i^z &= -\frac{\mu r_i^z}{\left(\left(\mathtt{r} + r_i^x\right)^2 + \left(r_i^y\right)^2 + \left(r_i^z\right)^2\right)^{\frac{3}{2}}} + \mathbf{P}_i^z + \mathbf{D}_i^z + u_i^z,
\end{aligned}$$

where $\dot{r}_i^x, \dot{r}_i^y, \dot{r}_i^z \in \mathbb{R}\left(\mathrm{m}\cdot\mathrm{s}^{-1}\right)$ are the relative velocities and $\ddot{r}_i^x, \ddot{r}_i^y, \ddot{r}_i^z \in \mathbb{R}\left(\mathrm{m}\cdot\mathrm{s}^{-2}\right)$ are the relative accelerations. Here, the orbital angular rate is $\tau \in \mathbb{R}\left(\mathrm{rad}\cdot\mathrm{s}^{-1}\right)$ with $\dot{\tau} = -2\dot{\mathtt{r}}\tau/\mathtt{r}$ and $\ddot{\mathtt{r}} = \mathtt{r}\tau^2 - \mu/\mathtt{r}^2$ from two-body relations with $\dot{\mathtt{r}}$ being the target's radial rate $\left(\mathrm{m}\cdot\mathrm{s}^{-1}\right)$. Non-control accelerations projected into the LOS frame are $\begin{bmatrix} p_i^\sigma & p_i^\gamma & p_i^\phi \end{bmatrix} \triangleq B^\top(q_i)\left(\mathbf{P}_i^{\mathrm{RTN}} + \mathbf{D}_i^{\mathrm{RTN}}\right) \in \mathbb{R}^3\left(\mathrm{m}\cdot\mathrm{s}^{-2}\right)$.

Substituting the Cartesian model and using the above kinematics yields the spherical dynamics

$$\ddot{q}_i = f_i(q_i, \dot{q}_i) + g(q_i)\, u_i + \omega(\tau, \dot{\tau}, q_i, \dot{q}_i),$$

with $f_i \triangleq \begin{bmatrix} f_i^{\sigma} & f_i^{\gamma} & f_i^{\phi} \end{bmatrix}^{\top} \in \mathbb{R}^3 \left(\mathrm{m}\cdot\mathrm{s}^{-2}, \mathrm{rad}\cdot\mathrm{s}^{-2}, \mathrm{rad}\cdot\mathrm{s}^{-2}\right)$ and
$\omega = \begin{bmatrix} \omega^{\sigma} & \omega^{\gamma} & \omega^{\phi} \end{bmatrix}^{\top} \in \mathbb{R}^3 \left(\mathrm{m}\cdot\mathrm{s}^{-2}, \mathrm{rad}\cdot\mathrm{s}^{-2}, \mathrm{rad}\cdot\mathrm{s}^{-2}\right)$ having components

$$
\begin{aligned}
f_i^{\sigma}\left(q_i, \dot{q}_i\right) &= \sigma_i \dot{\phi}_i^2 - \frac{\mu\left(\mathtt{r}\cos\gamma_i\cos\phi_i + \sigma_i\right)}{\left(\mathtt{r}^2 + \sigma_i^2 + 2\mathtt{r}\sigma_i\cos\gamma_i\cos\phi_i\right)^{\frac{3}{2}}} + \frac{\mu}{\mathtt{r}^2}\cos\gamma_i\cos\phi_i + p_i^{\sigma}, \\
f_i^{\gamma}\left(q_i, \dot{q}_i\right) &= \frac{\mu\left(\mathtt{r}\sin\gamma_i\sec\phi_i\right)}{\sigma_i\left(\mathtt{r}^2 + \sigma_i^2 + 2\mathtt{r}\sigma_i\cos\gamma_i\cos\phi_i\right)^{\frac{3}{2}}} + \frac{\mu}{\mathtt{r}^2\sigma_i}\sin\gamma_i\sec\phi_i + \frac{p_i^{\gamma}}{\sigma_i}, \\
f_i^{\phi}\left(q_i, \dot{q}_i\right) &= -\frac{2\dot{\sigma}_i\dot{\phi}_i}{\sigma_i} - \frac{\mu}{\mathtt{r}^2\sigma_i}\cos\gamma_i\sin\phi_i + \frac{\mu\mathtt{r}\cos\gamma_i\sin\phi_i}{\sigma_i\left(\mathtt{r}^2 + \sigma_i^2 + 2\mathtt{r}\sigma_i\cos\gamma_i\cos\phi_i\right)^{\frac{3}{2}}} + \frac{p_i^{\phi}}{\sigma_i},
\end{aligned}
$$

the input map $g\left(q_i\right) = \mathrm{diag}\left\{1, \frac{1}{\sigma_i}, \frac{1}{\sigma_i}\right\}$, and

$$
\begin{aligned}
\omega^{\sigma} &= \left(\tau + \dot{\gamma}_i\right)^2 \sigma_i \cos^2\phi_i, \\
\omega^{\gamma} &= 2\left(\tau + \dot{\gamma}_i\right)\dot{\phi}_i\tan\phi_i - \dot{\tau} - 2\left(\tau + \dot{\gamma}_i\right)\frac{\dot{\sigma}_i}{\sigma_i}, \\
\omega^{\phi} &= -\left(\tau + \dot{\gamma}_i\right)^2\sin\phi_i\cos\phi_i.
\end{aligned}
$$

The simulation enforces $\sigma_i \geq \sigma_{\min} > 0\,(\mathrm{m})$ (no collision or overlap), $|\phi_i| \leq \phi_{\max} < \pi/2\,(\mathrm{rad})$ (LOS avoids the polar singularity), and a target orbital radius $\mathtt{r} \in \left[r_p, r_a\right]\,(\mathrm{m})$. Together with $\eta \in \mathbb{T}^k$ and bounded initial relative velocities, all arguments of f_i, g, and ω remain in compact sets and the corresponding mappings are continuous.

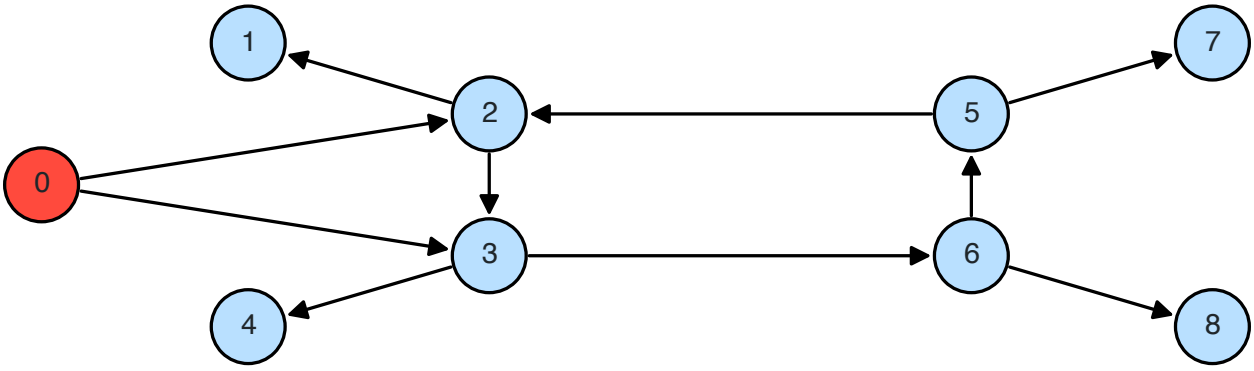

Figure 5-1. Communication topology visualization of the multi-spacecraft system.

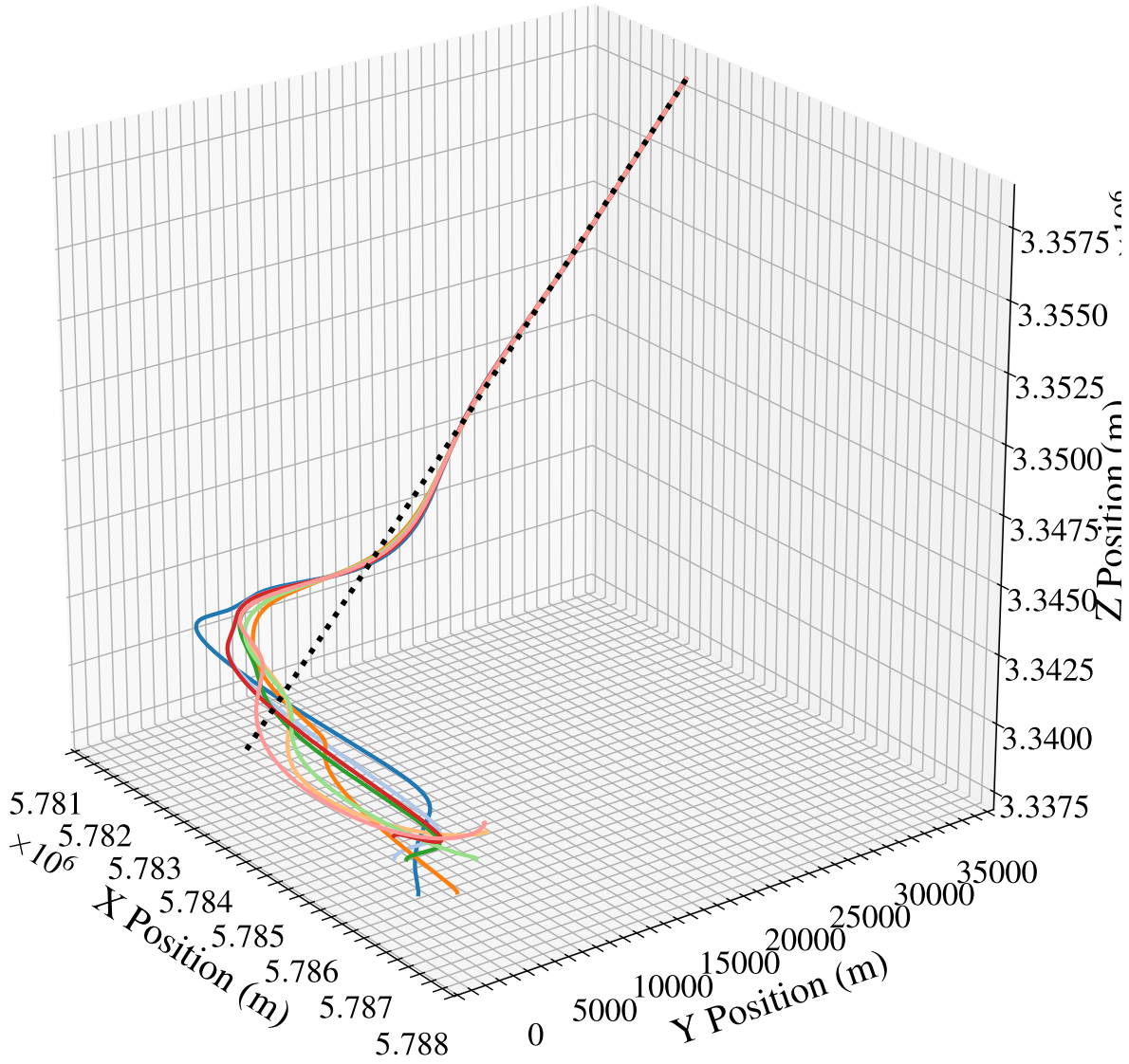

Figure 5-2. 3-D visualization of the trajectories servicing agents (solid lines) and the defunct spacecraft (dotted line) during the first five seconds of the simulation.

The defunct spacecraft is initialized on a near-Earth elliptical orbit with periapsis altitude 300×10^3 m, apoapsis altitude 700×10^3 m, and inclination $\pi/6$ rad. Earth's mean equatorial radius is taken as $R_E = 6,378 \times 10^3$ m. The corresponding radii are $r_p = R_E + 300 \times 10^3$ and $r_a = R_E + 700 \times 10^3$, given semi-major axis $a = \left(r_p + r_a\right)/2$ m. The periapsis orbital speed follows the vis-viva relation $v_p = \sqrt{\mu\left(2/r_p - 1/a\right)}\,\left(\mathrm{m} \cdot \mathrm{s}^{-1}\right)$.

Servicer initial positions are offset from the defunct spacecraft by random draws: radial distances are sampled uniformly from $U\,(2500, 5000)$ m, while azimuth γ and elevation ϕ offsets are sampled from $U\,(-0.5, 0.5)$ rad. The simulation commences in the Proximity Operations (Prox-Ops) phase, with each servicer initialized to the target's instantaneous inertial velocity; initial relative motion therefore arises solely from position offsets. The simulation time horizon is 130 seconds. The selected properties of all spacecraft are listed in Table 5-1.

Table 5-1. Spacecraft Simulation Parameters

Spacecraft Index	Mass (kg)	Cross-Sectional Area (m^2)
0	10,000.00	100.00
1	834.49	48.21
2	1079.90	19.78
3	914.07	39.79
4	828.70	26.91
5	649.89	28.83
6	977.69	46.35
7	670.44	4.48
8	1340.36	5.26

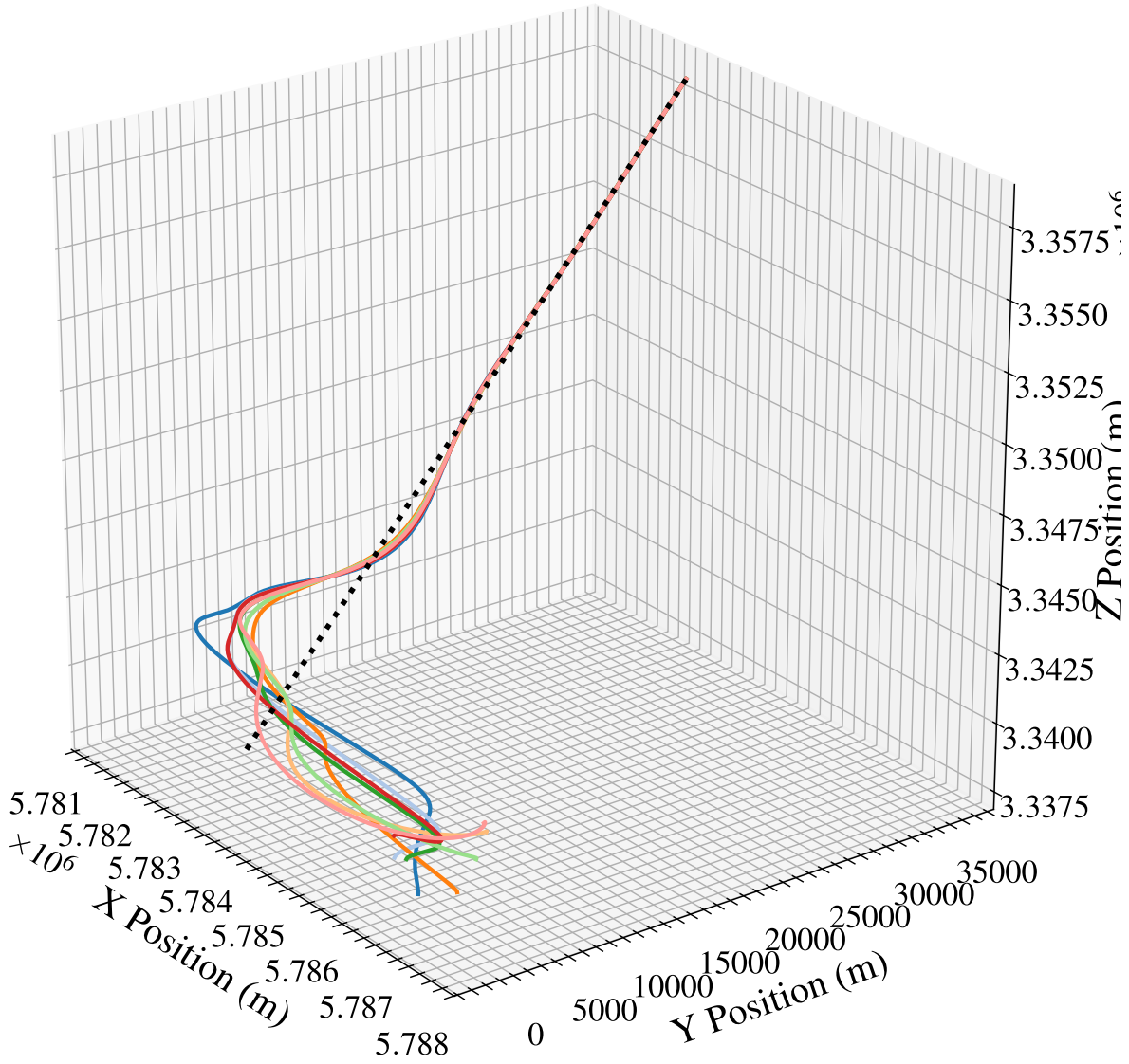

Figure 5-3. Plot showing the norm of the velocity estimation error, $\|\tilde{\zeta}_i\|$, over time for all agents $i \in \mathcal{V}$.

The multi-agent network uses the vertex set $\mathcal{V} = [8]$ and $\overline{\mathcal{V}} = \mathcal{V} \cup \{0\}$. The communication topology, depicted in Fig. 5-1, induces the in-Laplacian

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix},$$

with eigenvalues $\Lambda(L) = \{0, 1, 1 \pm i, 2\}$. The pinning matrix is given by $B = \text{diag}\{0, 1, 1, 0, 0, 0, 0, 0\}$, which specifies that agents 2 and 3 can sense the target state. A visualization of the communication topology is shown in Fig. 5-1.

Sensing is heterogeneous. Agent 2 carries a monocular optical camera providing angles-only LOS measurements to a noncooperative target, a modality demonstrated on the PRISMA/ARGON and AVANTI missions (e.g., [14]). For agent 2 ($i = 2$),

$$y_2 = \begin{bmatrix} \gamma_2 \\ \phi_2 \end{bmatrix} = C_2 q_2, \quad C_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Agent 3 uses a rendezvous radar that supplies range and azimuth, consistent with standard range-azimuth-elevation geometry. For agent 3 $i = 3$,

$$y_3 = \begin{bmatrix} \sigma_3 \\ \gamma_3 \end{bmatrix} = C_3 q_3, \quad C_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

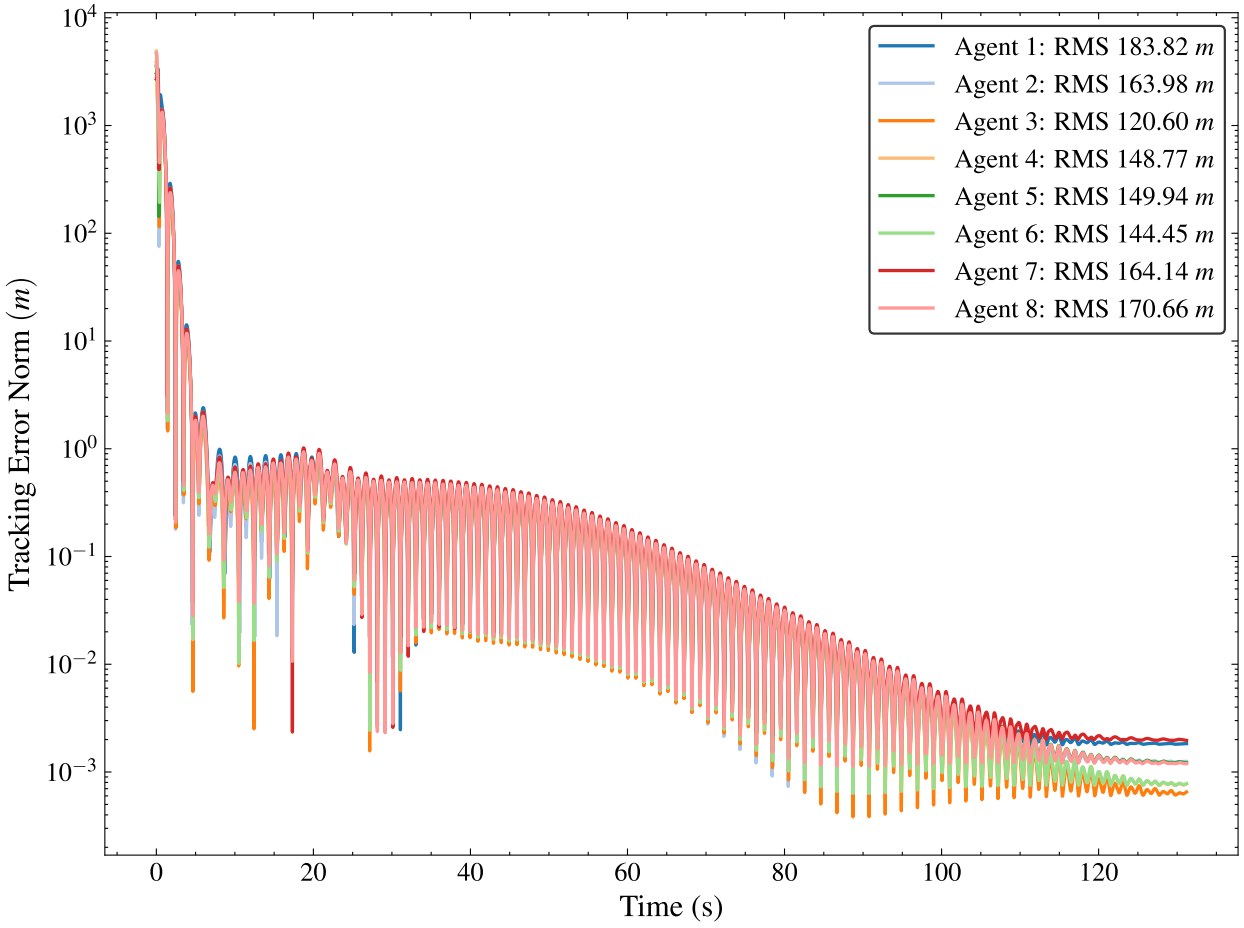

Figure 5-4. Plot showing the norm of the tracking error, $\|e_i\|$, over time for all agents $i \in \mathcal{V}$.

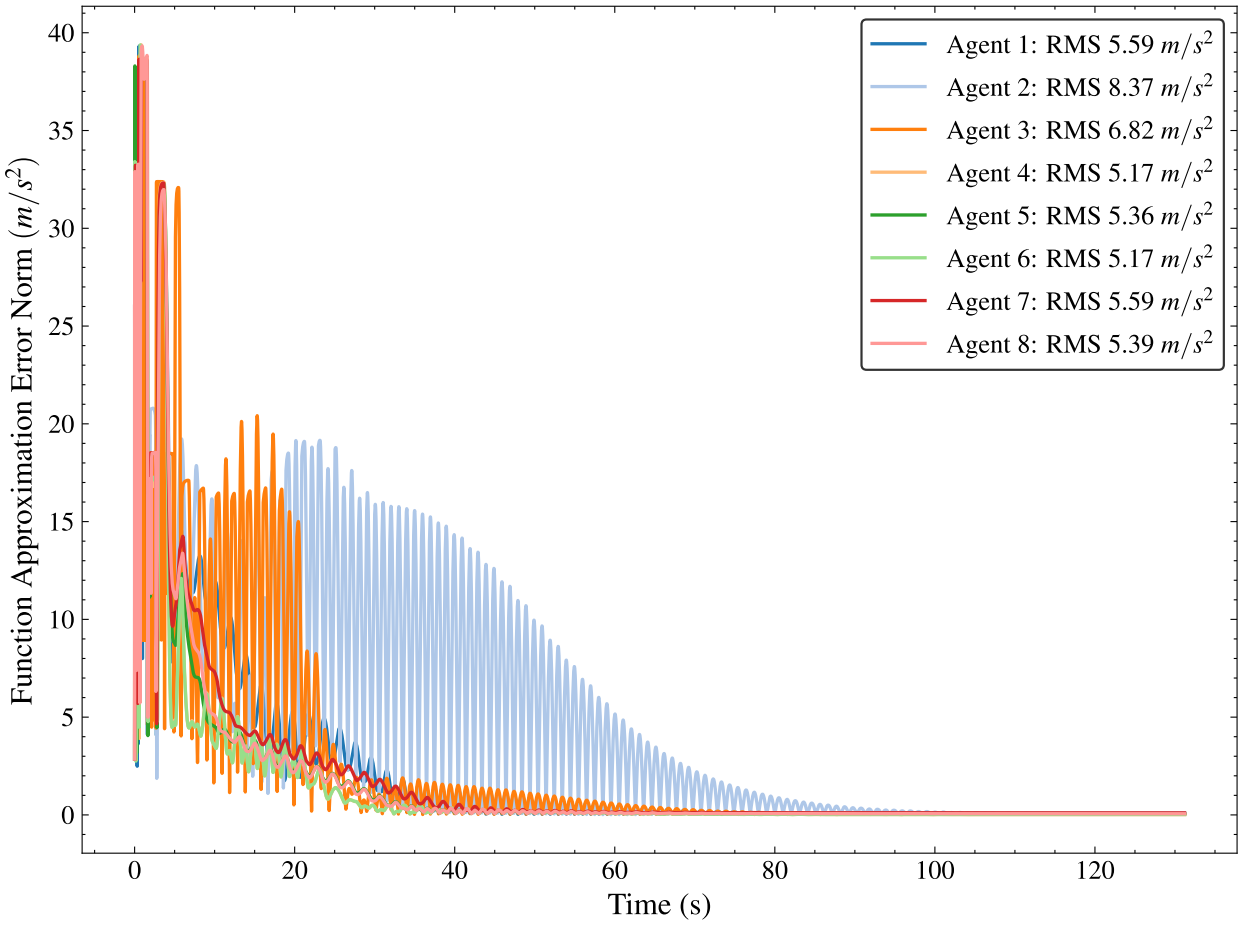

Figure 5-5. Plot showing the norm of the function approximation error, $\|\varepsilon_i(\kappa_i)\|$, over time for all agents $i \in \mathcal{V}$.

Each agent implements a residual network with four blocks and two hidden layers per block, using two neurons per hidden layer, for a total of 226 parameters. Hidden layers and the shortcut block in the ResNet use the swish activation function, and output layers use the tanh activation function. Gains are empirically selected as $k_1 = 10$, $k_2 = 15$, $k_3 = 2.5$, $k_4 = 10$, $k_5 = 5$, $k_6 = 0.0001$, and $\overline{\theta} = 25$, for all $i \in \mathcal{V}$. Fig. 5-2 shows the 3-D trajectories of all agents and the target, illustrating cooperative motion within the operational region.

Neighborhood velocity estimation performance appears in Fig. 5-3 as the time history of $\left\|\tilde{\zeta}_i\right\|$ for all $i \in \mathcal{V}$. All agents settle to approximate velocity estimation errors of 10.5 m/s after approximately 5 s. The pronounced initial velocity-estimation error results from zero-valued observer initialization (used in this simulation). When a priori velocity information is available, seeding the observer with those estimates, or with physics-based approximations, reduces the initial transient. Tracking performance appears in Fig. 5-4 as the time history of $\|e_i\|$ for all $i \in \mathcal{V}$. Agents achieve steady-state tracking errors near 1.0 m after approximately 5 s. The observed steady-state oscillations reflect the selected gains; reducing oscillations would increase control effort, indicating a practical trade-off between accuracy and energy. Fig. 5-5 reports the function approximation performance and demonstrates the capability of the ResNet architecture to approximate the unknown system dynamics, with all agents approaching a negligible function approximation error after approximately 100 seconds.

5.5 Conclusion

This work considered multi-agent target tracking for second-order agents over directed, rooted graphs using only relative position measurements and limited target information.

A trackability condition tailored to rooted digraphs was established, clarifying how information placement and graph structure determine feasibility. A locally implementable adaptation law was developed that relies solely on one-hop exchanges and does not require Laplacian symmetry. Unknown target dynamics were estimated online by a deep residual-network observer without offline training. A Lyapunov analysis guaranteed exponential convergence of the

target state estimation error and convergence of agent tracking errors to an explicit neighborhood, with bounds expressed in design parameters, graph quantities, and approximation accuracy.

A multi-spacecraft approach and servicing operation is simulated under representative directed topologies that corroborate the analysis and illustrate the closed-loop behavior, with all agents converging within 5 s to a neighborhood-velocity estimation error of approximately 10.5 m/s and a steady-state tracking error of approximately 1.0 m.

Future work will address more general output models, including state-dependent $C(x)$, time-dependent $C(t)$, jointly state- and time-dependent $C(x,t)$, and fully nonlinear measurement functions $y = h(x,t)$. Extending the trackability condition to these settings will require graph-aware observability notions that account for nonlinear and time-varying outputs.

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

This dissertation has developed a toolset for robust and adaptive target tracking in nonlinear multi-agent systems. The preceding chapters have systematically addressed the practical challenges of heterogeneity, model uncertainty, external disturbances, and constrained sensing and communication.

6.1 Conclusions

The primary contributions of this work provide assured, distributed control solutions for complex systems. Chapter 3 introduced an online ResNet adaptive controller for target tracking as well as a distributed observer to address the multi-agent target tracking problem under limiting sensing. Chapter 5 introduced a robust, RISE-based controller for heterogeneous, second-order nonlinear agents. By developing a novel P-function construction, this approach guarantees exponential tracking convergence in the presence of disturbances using only local, single-hop communication. Chapter 5 addressed the critical limitations of sensing and network topology by developing an adaptive control framework. This solution integrates a deep residual neural network observer to learn unknown target dynamics online, enabling assured tracking over directed graphs using only partial, relative state feedback.

The methods developed in this dissertation, while provably effective, are designed for specific problem classes. The control laws are tailored to the target tracking objective, and the stability analyses, while rigorous, are specific to the assumed system structures. This "case-by-case" design, which is common in the field, becomes difficult to scale when addressing a wider array of operational complexities, such as switching communication topologies, time-delays, heterogeneous objectives, or multi-target assignments. The remainder of this chapter outlines a path toward a unified, general framework for multi-agent coordination that is motivated by this limitation.

6.2 Future Work: Toward a Unified Framework for Multi-Agent Systems

The research presented in this dissertation can be extended and generalized along three primary axes: (1) expanding the complexity of the objective to include multi-target and formation

control, (2) generalizing the state space to formally handle heterogeneity, and (3) generalizing the network model to include spatiotemporal dynamics.

6.2.1 Immediate Generalizations: Multi-Target and Formation Objectives

The single-target tracking problem in Chapters 4 and 5 can be directly extended to a multi-target, multi-agent scenario. In this formulation, an assignment matrix B defines which of the N agents can sense which of the M targets. This defines a new disagreement signal η_i for each agent, which includes a sum of inter-agent errors (from the graph Laplacian L_A) and a sum of agent-target errors (from the assignment matrix).

The resulting ensemble disagreement η is a linear function of the tracking error e, $\eta = -\mathcal{H}e$, where the interaction matrix $\mathcal{H}$ takes the form $L_A + \Delta$. Here, Δ is a diagonal matrix derived from B that "pins" agents to their assigned targets. A controller based on this disagreement, $u_i = k\eta_i$, can be shown to be a gradient descent on a global quadratic cost function $\mathcal{L}(q)$. This demonstrates that the core concepts of this dissertation can be extended to more complex objectives. However, this approach still assumes agent homogeneity and becomes cumbersome when objectives are not simple relative positions.

6.2.2 A General Language: Cellular Sheaves for Heterogeneity

A more fundamental generalization is to redefine the system's mathematical structure using tools from algebraic topology. A cellular sheaf $\mathcal{F}$ over the agent graph G provides a formal language for handling heterogeneity.

Instead of assuming all agents live in $\mathbb{R}^n$, a sheaf assigns a unique vector space, or stalk $\mathcal{F}_i$, to each vertex i. This allows a system to natively include an $\mathbb{R}^3$ aerial vehicle and an $\mathbb{R}^2$ ground vehicle. Similarly, each edge ij is assigned its own stalk $\mathcal{F}_{ij}$, which acts as a "comparison space." The interaction between agents is defined by restriction maps $\mathcal{F}_{i \trianglelefteq ij}$, which are linear maps that project the agent states onto this common comparison space.

This framework elegantly recasts the control problem:

1. Heterogeneous Objectives: A UAV tracking a ground robot can be modeled by a restriction map that projects its $\mathbb{R}^3$ state onto the $\mathbb{R}^2$ edge space, while the ground robot's map is the identity.

2. Control as Consistency: The objective is to find a global section of the sheaf, which is a collection of agent states $\mathbf{q}$ that are mutually consistent on every edge (i.e., $\mathcal{F}_{i \trianglelefteq ij}(q_i) = \mathcal{F}_{j \trianglelefteq ij}(q_j)$.

3. The Sheaf Laplacian: The disagreement signal η used in this dissertation is revealed to be a special case of the sheaf Laplacian $L_{\mathcal{F}}$. The control objective $\eta = L_{\mathcal{F}} = 0$ is precisely the condition for $\mathbf{q}$ to be a global section.

This approach unifies consensus, formation control, and target tracking into the single, general problem of finding a global section in the kernel of the sheaf Laplacian.

6.2.3 Incorporating Spatiotemporal Dynamics

Finally, the static graph G must be replaced with a dynamic model to capture real-world communication. This includes switching topologies, time-delays, and environmental obstacles.

A rigorous approach is to model the network as a time-indexed family of graphs G_t. This graph is generated at each instant t by the intersection of "communication regions" $B_v(t)$, which are themselves functions of the agent's position, communication radius, and obstacles in the environment. Crucially, this family is not a filtration, as obstacles or agent movement can cause edges to disappear, breaking the assumption of monotone graph evolution.

To analyze such a system, the sheaf framework can be extended to a temporal sheaf. This is formally a functor from a time category into a category of network objects. This construction uses the categorical tools of pullbacks and pushouts to rigorously "glue" information across adjacent, and potentially asynchronous, time intervals. This provides a formal language for analyzing stability and performance in the presence of delays, packet loss, and dynamic, obstacle-aware network topologies.

LIST OF REFERENCES

[1] Aranda, M., López-Nicolás, G., Sagüés, C. and Zavlanos, M. M. [2016], 'Distributed formation stabilization using relative position measurements in local coordinates', *IEEE Trans. Autom. Control* **61**(12), 3925–3935.

[2] Bang, J. and Ahn, J. [2019], 'Multitarget rendezvous for active debris removal using multiple spacecraft', *J. Spacecr. Rockets* **56**(4), 1237–1247.

[3] Bidikli, B., Tatlicioglu, E. and Zergeroglu, E. [2014], A self tuning RISE controller formulation, *in* 'Proc. Am. Control Conf.', pp. 5608–5613.

[4] Boughellaba, M. and Tayebi, A. [2025], 'Distributed attitude estimation for multiagent systems on SO(3)', *IEEE Trans. Autom. Control* **70**(1), 657–664.

[5] Bullo, F. [2024], *Lectures on Network Systems*, 1.7 edn, Kindle Direct Publishing.

[6] Caughman, J. and Veerman, J. [2006], 'Kernels of directed graph Laplacians', *The electronic journal of combinatorics* **13**.

[7] Coddington, E. A., Levinson, N. and Teichmann, T. [1956], *Theory of ordinary differential equations*, American Institute of Physics.

[8] Dimarogonas, D. V., Frazzoli, E. and Johansson, K. H. [2012], 'Distributed event-triggered control for multi-agent systems', *IEEE Trans. Autom. Control* **57**, 1291–1297.

[9] Du, H., Cheng, Y., He, Y. and Jia, R. [2016], 'Second-order consensus for nonlinear leader-following multi-agent systems via dynamic output feedback control', *Int. J. Robust Nonlinear Control* **26**, 329–344.

[10] Eggleston, J. M. and Dunning, R. S. [1961], Analytical evaluation of a method of midcourse guidance for rendezvous with Earth satellites, NASA Tech. Note D-883, Hampton, VA.
URL: *https://ntrs.nasa.gov/citations/19980227200*

[11] Eugene Lavretsky, K. A. W. [2025], *Robust and Adaptive Control*, Springer Cham.

[12] Fang, X., Xie, L. and Li, X. [2024], 'Integrated relative-measurement-based network localization and formation maneuver control', *IEEE Trans. Autom. Control* **69**(3), 1906–1913.

[13] Fax, J. A. and Murray, R. M. [2004], 'Information flow and cooperative control of vehicle formations', *IEEE Trans. Autom. Control* **49**, 1465–1476.

[14] Gaias, G., Ardaens, J.-S. and Terzibaschian, T. [2015], Paving the way for future on-orbit-servicing missions: the AVANTI experiment, *in* 'Proc. 25th Int. Symp. Space Flight Dyn. (ISSFD)'.
URL: *https://elib.dlr.de/100930/*

[15] Goodfellow, I., Bengio, Y., Courville, A. and Bengio, Y. [2016], *Deep Learning*, Vol. 1, MIT press Cambridge.

[16] Hardt, M. and Ma, T. [2017], 'Identity matters in deep learning', *arXiv preprint arXiv.1611.04231* .

[17] He, K., Zhang, X., Ren, S. and Sun, J. [2016*a*], Deep residual learning for image recognition, *in* 'Proc. IEEE Conf. Comput. Vis. Pattern Recognit.', pp. 770–778.

[18] He, K., Zhang, X., Ren, S. and Sun, J. [2016*b*], 'Identity mappings in deep residual networks', *arXiv preprint arXiv.1603.05027* .

[19] Hong, Y., Chen, G. and Bushnell, L. [2008], 'Distributed observers design for leader-following control of multi-agent networks', *Automatica* **44**, 846–850.

[20] Hong, Y., Hu, J. and Gao, L. [2006], 'Tracking control for multi-agent consensus with an active leader and variable topology', *Automatica* **42**, 1177–1182.

[21] Horn, R. A. and Johnson, C. R. [1993], *Matrix Analysis*, Cambridge University Press., Cambridge.

[22] Hornik, K. [1991], 'Approximation capabilities of multilayer feedforward networks', *Neural Netw.* pp. 251–257.

[23] Hu, H.-X., Wen, G., Yu, X., Wu, Z.-G. and Huang, T. [2022], 'Distributed stabilization of heterogeneous mass in uncertain strong-weak competition networks', *IEEE Trans. Syst. Man Cybern. Syst.* **52**(3), 1755–1767.

[24] Hu, J. and Feng, G. [2010], 'Distributed tracking control of leader-follower multi-agent systems under noisy measurement', *Automatica* **46**(8), 1382–1387.

[25] Huang, G., Liu, Z., van der Maaten, L. and Weinberger, K. Q. [2017], Densely connected convolutional networks, *in* 'Proc. IEEE Conf. Comput. Vis. Pattern Recognit.'.

[26] Joshi, G. and Chowdhary, G. [2019], Deep model reference adaptive control, *in* 'Proc. IEEE Conf. Decis. Control', pp. 4601–4608.

[27] Kamalapurkar, R., Dixon, W. E. and Teel, A. [2020], 'On reduction of differential inclusions and Lyapunov stability', *ESAIM: Control, Optim. Calc. of Var.* **26**(24), 1–16.

[28] Kamalapurkar, R., Klotz, J. R., Walters, P. and Dixon, W. E. [2018], 'Model-based reinforcement learning in differential graphical games', *IEEE Trans. Control Netw. Syst.* **5**, 423–433.

[29] Kamalapurkar, R., Rosenfeld, J. A., Klotz, J., Downey, R. J. and Dixon, W. E. [2014], 'Supporting lemmas for RISE-based control methods', arXiv:1306.3432.

[30] Kawaguchi, K. and Bengio, Y. [2019], 'Depth with nonlinearity creates no bad local minima in ResNets', *Neural Netw.* **118**, 167–174.

[31] Kidger, P. and Lyons, T. [2020], Universal approximation with deep narrow networks, *in* 'Conf. Learn. Theory', pp. 2306–2327.

[32] Kim, H., Shim, H. and Seo, J. H. [2011], 'Output consensus of heterogeneous uncertain linear multi-agent systems', *IEEE Trans. Autom. Control* **56**, 200–206.

[33] Klotz, J. R., Kan, Z., Shea, J. M., Pasiliao, E. L. and Dixon, W. E. [2015], 'Asymptotic synchronization of a leader-follower network of uncertain Euler-Lagrange systems', *IEEE Trans. Control Netw. Syst.* **2**(2), 174–182.

[34] Klotz, J. R., Obuz, S., Kan, Z. and Dixon, W. E. [2018], 'Synchronization of uncertain euler–lagrange systems with uncertain time-varying communication delays', *IEEE Trans. Cybern.* **48**, 807–817.

[35] Kratsios, A. and Papon, L. [2022], 'Universal approximation theorems for differentiable geometric deep learning', *J. Mach. Learn. Res.* .

[36] Laczkovich, M. and Sós, V. T. [2015], *Real Analysis: Foundations and Functions of One Variable*, Undergraduate Texts in Mathematics, 1st edn, Springer, New York, NY. **URL:** *https://doi.org/10.1007/978-1-4939-2766-1*

[37] Lax, P. D. and Terrell, M. S. [2017], *Multivariable Calculus with Applications*, Springer International Publishing AG.

[38] Le, D., Greene, M., Makumi, W. and Dixon, W. E. [2022], 'Real-time modular deep neural network-based adaptive control of nonlinear systems', *IEEE Control Syst. Lett.* **6**, 476–481.

[39] Le, D., Patil, O., Nino, C. and Dixon, W. E. [2024], 'Accelerated gradient approach for deep neural network-based adaptive control of unknown nonlinear systems', *IEEE Trans. on Neural Netw. Learn. Syst.* .

[40] LeCun, Y., Bengio, Y. and Hinton, G. [2015], 'Deep learning', *Nature* **521**(436444).

[41] Lewis, F. L. [1999], 'Nonlinear network structures for feedback control', *Asian J. Control* **1**(4), 205–228.

[42] Lewis, F. L., Jagannathan, S. and Yesildirak, A. [1998], *Neural network control of robot manipulators and nonlinear systems*, CRC Press, Philadelphia, PA.

[43] Lewis, F. L., Zhang, H., Hengster-Movric, K. and Das, A. [2014], *Cooperative Control of Multi-Agent Systems: Optimal and Adaptive Design Approaches*, 1 edn, Springer London, London.

[44] Li, H., Xu, Z., Taylor, G., Studer, C. and Goldstein, T. [2018], Visualizing the loss landscape of neural nets, *in* 'Proc. 32nd Int. Conf. Neural Inf. Process. Syst.', Curran Assoc. Inc., pp. 6391–6401.

[45] Li, Y. and Tan, C. [2019], 'A survey of consensus for multi-agent systems', *Syst. Sci. Control Eng.* **7**, 468–482.

[46] Licitra, R. A., Bell, Z. I. and Dixon, W. E. [2019], 'Single-agent indirect herding of multiple targets with uncertain dynamics', *IEEE Trans. Robot.* **35**, 847–860.

[47] Lin, C., Lin, Z., Zheng, R., Yan, G. and Mao, G. [2016], 'Distributed source localization of multi-agent systems with bearing angle measurements', *IEEE Trans. Autom. Control* **61**, 1105–1110.

[48] Lin, H. and Jegelka, S. [2018], 'ResNet with one-neuron hidden layers is a universal approximator', *Adv. Neural Inf. Process. Syst.* **31**.

[49] Liu, C., Liang, E. and Chen, M. [2024], Characterizing ResNet's universal approximation capability, *in* 'Proc. Mach. Learn. Res.', Vol. 235, PMLR, pp. 31477–31515.

[50] Liu, M., Wan, Y., Lopez, V. G., Lewis, F. L., Hewer, G. A. and Estabridis, K. [2021], 'Differential graphical game with distributed global nash solution', *IEEE Trans. Control Netw. Syst.* **8**, 1371–1382.

[51] Loría, A. and Panteley, E. [2002], 'Uniform exponential stability of linear time-varying systems: revisited', *Syst. Control Lett.* **47**(1), 13–24.

[52] M Krstic, I Kanellakopoulos, P. K. [1995], *Nonlinear and Adaptive Control Design*, John Willey, New York.

[53] MacKunis, W., Patre, P., Kaiser, M. and Dixon, W. E. [2010], 'Asymptotic tracking for aircraft via robust and adaptive dynamic inversion methods', *IEEE Trans. Control Syst. Technol.* **18**(6), 1448–1456.
URL: *http://ncr.mae.ufl.edu/papers/CST10-2.pdf*

[54] Mei, J., Ren, W. and Chen, J. [2016], 'Distributed consensus of second-order multi-agent systems with heterogeneous unknown inertias and control gains under a directed graph', *IEEE Trans. Autom. Control* **61**(8), 2019–2034.

[55] Meng, X., Mei, J., Miao, Z., Wu, A. and Ma, G. [2024], 'Fully distributed consensus of multiple euler-lagrange systems under switching directed graphs using only position measurements', *IEEE Trans. Autom. Control* **69**, 1781–1788.

[56] Mesbahi, M. and Egerstedt, M. [2010], *Graph Theoretic Methods in Multiagent Networks*, Princeton Series in Applied Mathematics, Princeton Univ. Press, Princeton, NJ, USA.

[57] Nar, K. and Sastry, S. [2018], 'Residual networks: Lyapunov stability and convex decomposition', *arXiv preprint arXiv:1803.08203* .

[58] Nino, C. F., Patil, O. S. and Dixon, W. E. [2023], 'Second-order heterogeneous multi-agent target tracking without relative velocities', *IEEE Control Syst. Lett.* **7**, 3663–3668.

[59] Nino, C. F., Patil, O. S., Eisman, M. R. and Dixon, W. E. [2025], 'Online resnet-based adaptive control for nonlinear target tracking'.
URL: *https://arxiv.org/abs/2503.14372*

[60] Nino, C. F., Patil, O. S., Philor, J., Bell, Z. and Dixon, W. E. [2023], Deep adaptive indirect herding of multiple target agents with unknown interaction dynamics, *in* 'Proc. IEEE Conf. Decis. Control', pp. 2509–2514.

[61] Nino, C. F., Zegers, F. M., Phillips, S. and Dixon, W. E. [2023], Consensus over clustered networks using output feedback and asynchronous inter-cluster communication, *in* 'Am. Control Conf.', pp. 4844–4851.

[62] Nowzari, C., Garcia, E. and Cortés, J. [2019], 'Event-triggered communication and control of networked systems for multi-agent consensus', *Automatica* **105**, 1–27.

[63] Oh, K.-K., Park, M.-C. and Ahn, H.-S. [2015], 'A survey of multi-agent formation control', *Automatica* **53**, 424–440.

[64] Olfati-Saber, R., Fax, J. A. and Murray, R. M. [2007], 'Consensus and cooperation in networked multi-agent systems', *Proc. IEEE* **95**, 215–233.

[65] Olfati-Saber, R. and Murray, R. M. [2004], 'Consensus problems in networks of agents with switching topology and time-delays', *IEEE Trans. Autom. Control* **49**, 1520–1533.

[66] Papachristodoulou, A., Jadbabaie, A. and MÃŒnz, U. [2010], 'Effects of delay in multi-agent consensus and oscillator synchronization', *IEEE Trans. Autom. Cont.* **55**(6), 1471–1477.

[67] Patil, O., Isaly, A., Xian, B. and Dixon, W. E. [2022], 'Exponential stability with RISE controllers', *IEEE Control Syst. Lett.* **6**, 1592–1597.

[68] Patil, O., Le, D., Greene, M. and Dixon, W. E. [2022], 'Lyapunov-derived control and adaptive update laws for inner and outer layer weights of a deep neural network', *IEEE Control Syst Lett.* **6**, 1855–1860.

[69] Patil, O. S., Griffis, E. J., Makumi, W. A. and Dixon, W. E. [2023], 'Composite adaptive Lyapunov-based deep neural network (Lb-DNN) controller', *arXiv preprint arXiv.2311.13056* .

[70] Patil, O. S., Kamalapurkar, R. and Dixon, W. E. [to appear], 'Saturated rise controllers with exponential stability guarantees: A projected dynamical systems approach', *IEEE Trans. Autom. Control* .

[71] Patil, O. S., Le, D. M., Griffis, E. and Dixon, W. E. [2022], Deep residual neural network (ResNet)-based adaptive control: A Lyapunov-based approach, *in* 'Proc. IEEE Conf. Decis. Control', pp. 3487–3492.

[72] Patil, O. S., Stubbs, K., Amy, P. and Dixon, W. E. [2022], Exponential stability with RISE controllers for uncertain nonlinear systems with unknown time-varying state delays, *in* 'Proc. IEEE Conf. Decis. Control', pp. 6431–6435.

[73] Patre, P., Bhasin, S., Wilcox, Z. D. and Dixon, W. E. [2010], 'Composite adaptation for neural network-based controllers', *IEEE Trans. Autom. Control* **55**, 944–950.

[74] Penrose, R. [1955], 'A generalized inverse for matrices', *Math. Proc. Camb. Philos. Soc.* **51**(3), 406–413.

[75] Petros A. Ioannou, J. S. [2012], *Robust Adaptive Control*, Dover Books on Electrical Engineering Series.

[76] Pettersen, K. Y. [2017], 'Lyapunov sufficient conditions for uniform semiglobal exponential stability', *Automatica* **78**, 97–102.

[77] Phillips, S. and Sanfelice, R. G. [2019], 'Robust distributed synchronization of networked linear systems with intermittent information', *Automatica* **105**, 323–333.

[78] Qin, J., Ma, Q., Shi, Y. and Wang, L. [2017], 'Recent advances in consensus of multi-agent systems: A brief survey', *IEEE Trans. Ind. Electron.* **64**, 4972–4983.

[79] Qin, J. and Yu, C. [2013], 'Cluster consensus control of generic linear multi-agent systems under directed topology with acyclic partition', *Automatica* **49**, 2898–2905.

[80] Qu, Z. [2009], *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*, 1 edn, Springer, London, U.K.

[81] Ren, W. and Cao, Y. [2013], *Distributed Coordination of Multi-agent Networks*, Springer London.

[82] Rolnick, D. and Tegmark, M. [2018], The power of deeper networks for expressing natural functions, *in* 'Int. Conf. Learn. Represent.'.

[83] Ryu, M. and Choi, K. [2024], 'CNN-based end-to-end adaptive controller with stability guarantees', *arXiv preprint arXiv.2403.03499* .

[84] Saberi, A., Stoorvogel, A. A., Zhang, M. and Sannuti, P. [2022], *Synchronization of Multi-Agent Systems in the Presence of Disturbances and Delays*, Birkhauser Cham.

[85] Sakurama, K. [2021], 'Unified formulation of multiagent coordination with relative measurements', *IEEE Trans. Autom. Control* **66**, 4101–4116.

[86] Sastry, S. and Bodson, M. [2011], *Adaptive Control: Stability, Convergence, and Robustness*, Dover.

[87] Schaub, H. and Junkins, J. [2003], *Analytical Mechanics of Space Systems*, AIAA Education Series, New York.

[88] Sebastián, E., Montijano, E. and Sagüés, C. [2022], 'Adaptive multirobot implicit control of heterogeneous herds', *IEEE Trans. Robot.* **38**, 3622–3635.

[89] Semsar-Kazerooni, E. and Khorasani, K. [2008], 'Optimal consensus algorithms for cooperative team of agents subject to partial information', *Automatica* **44**, 2766–2777.

[90] Shevitz, D. and Paden, B. [1994], 'Lyapunov stability theory of nonsmooth systems', *IEEE Trans. Autom. Control* **39 no. 9**, 1910–1914.

[91] Stone, M. H. [1948], 'The generalized Weierstrass approximation theorem', *Math. Mag.* **21**(4), 167–184.

[92] Sun, C., Ye, M. and Hu, G. [2017], 'Distributed time-varying quadratic optimization for multiple agents under undirected graphs', *IEEE Trans. Autom. Control* **62**(7), 3687–3694.

[93] Sun, R., Greene, M., Le, D., Bell, Z., Chowdhary, G. and Dixon, W. E. [2022], 'Lyapunov-based real-time and iterative adjustment of deep neural networks', *IEEE Control Syst. Lett.* **6**, 193–198.

[94] Tabuada, P. and Gharesifard, B. [2023], 'Universal approximation power of deep residual neural networks through the lens of control', *IEEE Trans. Autom. Control* **68**(5), 2715–2728.

[95] Teo, J., How, J. P. and Lavretsky, E. [2009], On approximate dynamic inversion and proportional-integral control, *in* 'Proc. Am. Control. Conf.', pp. 1592–1597.

[96] Vamvoudakis, K. G., Lewis, F. L. and Hudas, G. R. [2012], 'Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality', *Automatica* **48**, 1598–1611.

[97] Veit, A., Wilber, M. and Belongie, S. [2016], Residual networks behave like ensembles of relatively shallow networks, *in* 'Proc. 30th Int. Conf. Neural Inf. Process. Syst.', Curran Assoc. Inc., pp. 550–558.

[98] Wang, X., Sun, J., Wu, Z. and Li, Z. [2022], 'Robust integral of sign of error-based distributed flocking control of double-integrator multi-agent systems with a varying virtual leader', *Int. J. Robust Nonlinear Control* **32**(1), 286–303.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/rnc.5823*

[99] Yang, Y., Mei, J., Shi, X. and Ma, G. [2025], 'Fully distributed consensus of multiple euler-lagrange systems with time-varying asymmetric full-state constraints', *IEEE Trans. on Autom. Control* **70**(8), 5483–5490.

[100] Zegers, F. M., Deptula, P., Shea, J. M. and Dixon, W. E. [2022], 'Event/self-triggered approximate leader-follower consensus with resilience to byzantine adversaries', *IEEE Trans. Autom. Cont.* **67**(3), 1356–1370.

[101] Zhang, X.-R., Zhao, Z.-L. and Liu, T. [2025], 'Finite-time consensus for high-order multi-agent systems under directed communication topologies', *IEEE Trans. on Autom. Control* pp. 1–8.

BIOGRAPHICAL SKETCH

Dr. Cristian F. Nino received his Ph.D. in mechanical engineering from the University of Florida in 2025, where he was advised by Dr. Warren E. Dixon. He earned an M.S. in mechanical engineering (control systems) in 2024 and dual B.S. degrees in mechanical engineering (design and manufacturing) and mathematics in 2021, all from the University of Florida. His research focuses on robust adaptive nonlinear control and multi-agent systems, with applications to target tracking, distributed state and parameter estimation, and guidance, navigation, and control. He also studies learning-theoretic tools for safety-critical assured autonomy, including reinforcement learning, Lyapunov-based deep learning, and spatiotemporal graph models. His work has been supported by the U.S. Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship, the National Science Foundation Research in Undergraduate Institutions (RUI) program, and the Machen Florida Opportunity Scholarship, and he is the recipient of the University of Florida Graduate Student Research Award (2024-2025).