

**ARTIFICIAL INTELLIGENCE-BASED SECURE FORMATION
MANEUVER CONTROL OF MULTI-AGENT CYBER-PHYSICAL SYSTEMS
WITH GUARANTEED PERFORMANCE**

BY

MUHAMMD MAARUF

A Dissertation Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In

SYSTEMS AND CONTROL ENGINEERING

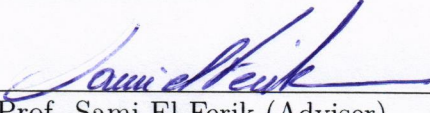
OCTOBER 2023

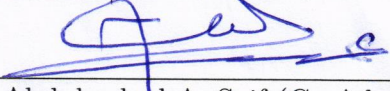
KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **MUHAMMAD MAARUF** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY IN CONTROL AND INSTRUMENTATION ENGINEERING**.

Thesis Committee


Prof. Sami El-Ferik (Adviser)


Prof. Abdulwahed A. Saif (Co-Advisor)


Dr. Mujahed M. Aldheefallah (Member)


Dr. Muhammad F. Mysorewala (Member)


Dr. Muhammad Khalid (Member)


Dr. Mujahed M. Aldheefallah
Department Chairman


Prof. Suliman S. Al-Homidan
Dean of Graduate Studies

Date



©Muhammad Maaruf
2023

To my parents and brothers

ACKNOWLEDGMENTS

In the name of Allah (SWT), the Most Gracious, the Most Merciful, and whose blessings are endless. First and foremost I would like to appreciate the Kingdom of Saudi Arabia for sponsoring my Ph.D. program at King Fahd University of Petroleum and Minerals. I would like to express my gratitude for the supervision and guidance I obtained from my supervisor, Professor Sami El-Ferik. He helped me create the thesis topic and main objectives that distinguished it from existing works. Moreover, I would like to thank my thesis committee members for sparing their time to review my research and offer useful advice which helped in improving the research.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	vi
LIST OF TABLES	xii
LIST OF FIGURES	xiii
ABSTRACT (ENGLISH)	xvi
ABSTRACT (ARABIC)	xx
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.1.1 Multi-agent cyber-physical systems	1
1.1.2 Multi-agent formation control:	3
1.1.3 Multi-agent affine Formation Control	5
1.1.4 Security challenge of multi-agent cyber-physical systems	6
1.1.5 Artificial intelligence in multi-agent cyber-physical systems	7
1.2 Motivation	8
1.3 Objectives of the Thesis	9
1.4 Thesis Structure and Contributions	10
1.5 Publications	14
CHAPTER 2 PRELIMINARIES	16
2.1 Notations	16
2.2 Graph Theory	16

2.3	Formation Graph	17
2.4	Affine Formation Transformation	17
2.4.1	Stress matrix	18
CHAPTER 3 LITERATURE REVIEW		21
3.1	Existing Literature on affine formation maneuver control	21
3.1.1	Affine formation maneuver control of MASs described by inte- grator dynamic equations	22
3.1.2	Affine formation maneuver control of MASs described by gen- eral linear dynamic equations	24
3.1.3	Affine formation maneuver control of MASs described by non- linear dynamic equations	25
3.2	Research gaps	27
3.3	Proposed control problem	28
CHAPTER 4 REINFORCEMENT LEARNING-BASED CONTROL STRATEGY FOR MULTI-AGENT SYSTEMS SUBJECTED TO ACTUATOR CYBERATTACKS DURING AFFINE FORMATION MANEUVERS		29
4.1	Introduction	30
4.2	Problem formulation	32
4.2.1	Dynamic models of the agents	32
4.2.2	Neural network approximations	33
4.2.3	Actuator attacks	34
4.3	Control design	36
4.3.1	Critic design	37
4.3.2	Actor design	40
4.4	Simulation results	51
CHAPTER 5 SECURE REINFORCEMENT LEARNING-BASED FINITE-TIME AFFINE FORMATION MANEUVER CONTROL		

OF MULTI-AGENT SYSTEMS SUBJECTED TO DECEPTION ATTACKS AND BACKLASH-LIKE HYSTERESIS 59

5.1	Introduction	60
5.2	Problem Formulation	62
5.2.1	Dynamic models of the agents	62
5.2.2	Backlash-like Hysteresis	62
5.2.3	Sensor Attack model	63
5.3	Control design	65
5.3.1	Critic design	67
5.3.2	Actor design	69
5.4	Simulation Results	79

CHAPTER 6 SECURE FAULT-TOLERANT REINFORCEMENT LEARNING-BASED PRESCRIBED PERFORMANCE AFFINE FORMATION MANEUVER CONTROL OF MULTI-AGENT SYS- TEMS WITH SENSOR DECEPTION ATTACK 88

6.1	Introduction	89
6.2	Problem Formulation	92
6.2.1	Dynamic models of the agents	93
6.2.2	Neural network approximations	96
6.2.3	Finite Time Prescribed Performance Function	97
6.3	Control design	102
6.4	Simulation results	123

CHAPTER 7 SAFETY-GUARANTEED SECURE REINFORCE- MENT LEARNING-BASED AFFINE FORMATION MANEUVER CONTROL OF QUADROTORS IN ADVERSARIAL ENVIRON- MENTS 134

7.1	Introduction	135
7.2	Problem Formulation	139
7.2.1	Quadrotor Dynamic Model	139

7.2.2	Radial Basis Function Neural Network	145
7.2.3	Barrier Function	146
7.3	Control Design	146
7.3.1	Critic design	150
7.3.2	Actor design	152
7.4	Simulation results	160
 CHAPTER 8 NEURAL NETWORK-BASED CONTROL FOR AFFINE FORMATION MANEUVER OF MULTI-AGENT SYS- TEMS WITH EXTERNAL DISTURBANCES		 166
8.1	Introduction	167
8.2	Problem formulation	167
8.2.1	Neural network	168
8.3	Control design	169
8.4	Simulation results	173
 CHAPTER 9 CRITIC LEARNING-BASED ZERO-SUM DIFFER- ENTIAL GAME FOR AFFINE FORMATION MANEUVER OF MULTI-AGENT SYSTEMS WITH ACTUATOR ATTACKS		 178
9.1	Introduction	179
9.2	Problem formulation	183
9.2.1	Dynamics of the agents	183
9.3	Control Design	185
9.3.1	Approximate solution of the HJI Equation Using Critic Learn- ing Mechanism	188
9.3.2	Stability analysis of the closed-loop system	192
9.4	Simulation results	198
 CHAPTER 10 CONCLUSIONS AND FUTURE WORK		 206
10.1	Conclusions	206
10.2	Future Works	209

REFERENCES	211
VITAE	247

LIST OF TABLES

9.1 Comparison of the mean-squared tracking errors of the proposed control technique and that in (El-Ferik et al 2023). 205

LIST OF FIGURES

2.1	Various affine transformations of a nominal configuration.	18
4.1	The nominal formation showing the inter-agents communications. . . .	51
4.2	Effects of attacks on the trajectories of the leader-follower MAS under the control input proposed in (Zhi and Chen 2021).	54
4.3	The trajectories of the leader-follower MAS under the attack mitigation control protocol (4.34).	55
4.4	The AFM tracking errors between the followers and the leaders under the proposed scheme.	55
4.5	The RL control signals of the followers that mitigate actuator attacks and maintain the AFM.	56
4.6	The update of the actor weights for the RBFNN approximation of $f(q, v)$.	56
4.7	The update of the actor weights for the RBFNN approximation of $\delta(q, v)$.	57
4.8	The update of the critic weights for the RBFNN approximation of $I(t)$.	57
4.9	The approximation errors between $f(q, v)$ and $\hat{f}(q, v)$	58
4.10	Trajectories showing various affine formation maneuvers of the leaders and followers.	58
5.1	The nominal formation graph of the agents.	81
5.2	The finite-time secure control protocols	83
5.3	The tracking errors of the followers.	84
5.4	The norm-2 of weight vector of the actor neural network for estimating F	85

5.5	The norm-2 of weight vector of the actor neural network for estimating G	86
5.6	The norm-2 of weight vector of the critic neural network.	87
5.7	Various AFMs of the leader-follower MAS	87
6.1	A simple example showing the performance of the proposed performance function.	100
6.2	A simple example showing the performance of existing performance functions.	102
6.3	The leader-follower nominal formation topology.	123
6.4	Optimized virtual control protocols of the four followers.	127
6.5	Optimized actual control protocols of the four followers.	128
6.6	Time-varying trajectories of the followers and the leaders with control method.	129
6.7	Prescribed leader-follower tracking errors.	130
6.8	Various leader-follower affine formation maneuvering.	131
6.9	Time-varying trajectories of the followers and the leaders with control method in (Zhi et al. 2021).	132
6.10	Various leader-follower affine formation maneuvering with control method in (Zhi et al. 2021).	133
7.1	The nominal formation of the quadrotors.	160
7.2	Leader-follower Time-varying positions trajectories.	163
7.3	Tracking errors of the followers.	163
7.4	Actor-critic learning control protocols of the followers.	164
7.5	Norm-2 of the actor weights.	164
7.6	Norm-2 of the critic weights.	165
7.7	Leader-follower affine formation maneuvers.	165
8.1	Undirected communication graph showing inter-agent communication. .	173
8.2	Two-dimensional affine formation maneuvering of the seven quadrotors.	177

9.1	Target formation depicting the communication pattern and positions of the seven agents.	198
9.2	Evolution of the norm-2 of the Weights of the critic networks of the agents.	201
9.3	The attack-tolerant optimal control policies.	201
9.4	The tracking errors of the followers.	202
9.5	The affine transformations of the leader-follower nominal formation of the seven agents.	203
9.6	The affine formation maneuvers of the seven agents under the control approach in (Zhi et al. 2021).	203
9.7	The tracking errors of the followers under the AFM control strategy proposed in (El-Ferik et al 2023).	204

THESIS ABSTRACT

NAME: Muhammad Maaruf

TITLE OF STUDY: Artificial Intelligence-based Secure Formation Maneuver
Control of multi-agent Cyber-Physical Systems with guaranteed performance

MAJOR FIELD: Control and Instrumentation Engineering

DATE OF DEGREE: October 2023

Affine formation maneuver control of leader-follower multi-agent systems is a new research field that is yet to be fully explored. In this approach, only the leaders have information about the formation maneuvers. The affine formation maneuvers of the followers is uniquely determined by the position of the leaders. Several formation maneuvers such as translation, rotation, scaling, shearing, and collinearity can be realized simultaneously with the aid of stress matrices and the principles of affine transformation. Multi-agent systems are interacting via a cyber-physical framework where information transmission is facilitated to maintain the target formation and avoid collisions. This wireless interaction exposes the multi-agent cyber-physical systems to malicious cyber-attacks. These attacks mostly target the systems' sensors,

actuators, or communication channels, which may lead to the deterioration of the acquired formation, and consequently lead to the failure of the entire system. Therefore, guaranteeing the security and performance of multi-agent cyber-physical systems during affine formation maneuvers is critical. This thesis proposes appropriate control and learning based on artificial Intelligence to counter the threats imposed by cyber-attacks.

The first issue investigated in this thesis is the reinforcement learning-based control strategy for second-order continuous-time multi-agent systems subjected to actuator cyber-attacks during affine formation maneuvers. In this case, a long-term performance index is created to track the multi-agent systems' tracking faults using a leader-follower structure. To approximate the ideal solution, which is challenging to find for systems vulnerable to cyber-attacks during time-varying maneuvers, a critical neural network is used. The distributed control protocol is obtained, and the long-term performance index is minimized, using an actor neural network strengthened with critic signals. The actor-critic neural networks calculate unknown dynamics and the severity of attacks on the multi-agent systems' actuators. The Nussbaum functions are applied to address this issue since attacks can result in a loss of control direction.

The second issue investigates the fast finite-time reinforcement learning-based control method for secure affine formation maneuvers of second-order nonlinear multi-agent systems subjected to sensor deception attacks and backlash-like hysteresis. The common assumptions that the input gains are known, the slope of the backlash-like hysteresis is bounded by known constants, and the time-varying gains induced by the deception attacks are bounded by known constants are eliminated. Through the theory of finite-

time convergence and fast terminal sliding mode surface, the actor-critic reinforcement learning architecture is employed to steer the followers to the target formation maneuvers, counter the sensor deception attacks, and tackle the actuator backlash-like hysteresis without any prior knowledge about them.

The third issue is the design of an optimized secure backstepping control with prescribed performance for affine formation maneuver of nonlinear MAS with unknown deception attacks and actuator faults. The control objective, in this case, is to ensure that the leader-follower maneuvers are resilient in the event of deception attacks and actuator faults. Moreover, prescribed performance functions are deployed to make sure the tracking errors are within a predefined set during the attacks. The optimal backstepping control laws are obtained from the solution of the Hamilton-Jacobi-Bellman equation. Nonetheless, the solution to the equation is difficult to obtain due to the nonlinearity of the multi-agent systems, and the deception attacks. This problem is tackled by employing a reinforcement learning approach using the identifier-actor-critic architecture.

The fourth issue aims to study the secure and safe leader-follower affine formation maneuver of a group of quadrotors under external disturbances, sensor deception attacks, and actuator injection attacks. To ensure the safe operation of the quadrotors considering their physical limitations, their dynamic states are constrained within a safe region of operation. To ensure that the safety constraints are not violated, a barrier Lyapunov function is employed. Then, the collective maneuvering of the quadrotors and their resilience against cyber-attacks are guaranteed using a new distributed sliding mode control and reinforcement learning.

The fifth issue is the anti-attack affine formation maneuver control of nonlinear multi-agent systems subjected to adversarial disturbances. Here, an artificial neural network is employed as an intelligent framework to detect, estimate, and mitigate harmful cyber-attacks. Then, a nonlinear control technique based on the variable structure is utilized to achieve a resilient leader-follower collective formation maneuvering despite the attacks. The proposed control algorithm is applied to a system of multiple quadrotor UAVs. The quadrotors maintained their collective formation maneuvering in the presence of adversarial disturbances.

The sixth issue proposes using the zero-sum differential game algorithm to address the affine formation maneuver control problem of a second-order multi-agent system subjected to actuator attacks. In this scheme, the distrusted optimal control policy and the actuator attack signals are treated as two opposing teams. The leader-follower distributed optimal control policies and the cyber-attack attenuation policies are derived from the Nash equilibrium solution of the Hamilton-Jacobi-Isaac (HJI) equation. The solution of the HJI equation is estimated with the aid of the critic neural network. The weights of the critic system are learned online based on the adaptive dynamic programming (ADP) scheme.

ملخص الرسالة

الاسم:

محمد معروف

عنوان الدراسة:

التحكم في مناورة التشكيل الآمن القائم على الذكاء الاصطناعي للأنظمة متعددة الوكلاء

التخصص:

هندسة النظم والتحكم

تاريخ الدرجة العلمية:

أكتوبر ٢٠٢٣

يعد التحكم في تكوين الأنظمة مجالاً بحثياً جديداً لم يتم التحقيق فيه بشكل كامل. القادة فقط لديهم معلومات حول المناورات. يتم تحديد تشكيل الأتباع بشكل فريد من خلال موقف القادة. يمكن تحقيق العديد من التشكيلات مثل الترجمة، والتدوير، والقياس، والقص، والعلاقة الخطية المتداخلة في وقت واحد بمساعدة مصفوفات الإجهاد ومبادئ التحول المتقارب. تتفاعل الأنظمة متعددة الوكلاء عبر إطار مادي إلكتروني حيث يتم تسهيل نقل المعلومات للحفاظ على تكوين الهدف وتجنب الاصطدامات. يعرض هذا التفاعل اللاسلكي الأنظمة المادية السيبرانية متعددة الوكلاء للهجمات السيبرانية الضارة. تستهدف هذه الهجمات في الغالب مستشعرات النظام أو مشغلاته أو قنوات الاتصال، مما قد يؤدي إلى تدهور التكوين المكتسب، وبالتالي يؤدي إلى فشل النظام بأكمله. لذلك، يعد ضمان أمان وأداء الأنظمة المادية السيبرانية متعددة الوكلاء أثناء مناورات التشكيل المتقارب أمراً بالغ الأهمية. تقترح هذه الأطروحة التحكم المناسب والتعلم القائم على الذكاء الاصطناعي لمواجهة التهديدات التي تفرضها الهجمات السيبرانية.

القضية الأولى التي تم بحثها في هذه الأطروحة هي استراتيجية التحكم المبنية على التعلم المعزز للأنظمة متعددة الوكلاء من الدرجة الثانية والمستمرة والتي تتعرض لهجمات إلكترونية مشغلة أثناء مناورات التشكيل المتقارب. في هذه الحالة، يتم إنشاء مؤشر أداء طويل المدى لتتبع أخطاء تتبع الأنظمة متعددة الوكلاء باستخدام هيكل القائد-التابع. من أجل تقريب الحل

المثالي، والذي يصعب العثور عليه للأنظمة المعرضة للهجمات السيبرانية أثناء المناورات المتغيرة بمرور الوقت، يتم استخدام شبكة عصبية مهمة. يتم الحصول على بروتوكول التحكم الموزع، ويتم تقليل مؤشر الأداء على المدى الطويل، باستخدام شبكة عصبية فاعلة معززة بالإشارات الناقدة. تقوم الشبكات العصبية الناقدة للممثل بحساب الديناميكيات غير المعروفة وشدة الهجمات على مشغلات الأنظمة متعددة الوكلاء. يتم تطبيق وظائف muabssuN لمعالجة هذه المشكلة نظرًا لأن الهجمات يمكن أن تؤدي إلى فقدان اتجاه التحكم.

تبحث المشكلة الثانية في طريقة التحكم السريعة القائمة على التعلم المعزز في الوقت المحدد لمناورات التشكيل المتقارب الآمن للأنظمة متعددة الوكلاء غير الخطية من الدرجة الثانية المعرضة لهجمات خداع الاستشعار والتباطؤ الشبيه برد الفعل العكسي. الافتراضات الشائعة بأن مكاسب المدخلات معروفة، وأن ميل التباطؤ الشبيه برد الفعل العكسي يحده ثوابت معروفة، ويتم التخلص من المكاسب المتغيرة بمرور الوقت الناجمة عن هجمات الخداع بواسطة ثوابت معروفة. من خلال نظرية التقارب في الوقت المحدد وسطح وضع الانزلاق الطرقي السريع، يتم استخدام بنية التعلم المعزز للناقد والممثل لتوجيه المتابعين إلى مناورات تشكيل الهدف، ومواجهة هجمات خداع المستشعر، ومعالجة التباطؤ الذي يشبه رد الفعل العكسي للمشغل دون أي معرفة مسبقة عنهم.

المسألة الثالثة هي تصميم نظام التحكم الخلفي الآمن الأمثل مع الأداء الموصوف لمناورة التشكيل المتقارب ل SAM غير الخطية مع هجمات الخداع غير المعروفة وأخطاء المشغل. هدف التحكم، في هذه الحالة، هو التأكد من أن مناورات القائد والأنواع مرنة في حالة وقوع هجمات الخداع وأخطاء المشغل. علاوة على ذلك، يتم نشر وظائف الأداء المحددة للتأكد من أن أخطاء التتبع تقع ضمن مجموعة محددة مسبقًا أثناء الهجمات. تم الحصول على قوانين التحكم الأمثل للخلف من حل معادلة هاملتون-جاكوبي-بيلمان. ومع ذلك، يصعب الحصول على حل للمعادلة بسبب عدم خطية الأنظمة متعددة الوكلاء، وهجمات الخداع. تتم معالجة هذه المشكلة من خلال استخدام نهج التعلم المعزز باستخدام بنية المعرفة-الممثل-الناقد. أما المبحث الرابع فيهدف إلى دراسة مناورة التشكيل التقاربي الآمن والتابع لمجموعة من المحركات الرباعية تحت تأثير الاضطرابات الخارجية وهجمات الخداع الاستشعاري وهجمات حقن المشغل. لضمان التشغيل الآمن للمحركات الرباعية مع الأخذ في الاعتبار حدودها المادية، تكون حالاتها الديناميكية مقيدة داخل منطقة تشغيل آمنة. ولضمان عدم انتهاك قيود السلامة، يتم استخدام وظيفة حاجز vonupayL. ومن ثم، المناورة الجماعية للطائرات الرباعية وقدرتها على الصمود في مواجهة الهجمات السيبرانية

المسألة الخامسة هي التحكم في مناورة التشكيل المتقارب المضاد للهجوم للأنظمة غير الخطية متعددة الوكلاء المعرضة لاضطرابات الخصومة. هنا، يتم استخدام الشبكة العصبية الاصطناعية كإطار ذكي لاكتشاف الهجمات الإلكترونية الضارة وتقديرها والتخفيف من حدتها. بعد ذلك، تم استخدام تقنية تحكم غير خطية تعتمد على البنية المتغيرة لتحقيق مناورة تشكيلية جماعية مرنة بين القائد والأتباع بالرغم من الهجمات. يتم تطبيق خوارزمية التحكم المقترحة على نظام متعدد الطائرات بدون طيار رباعية المحركات. حافظت المحركات الرباعية على مناوراتها الجماعية في ظل وجود اضطرابات معادية. يقترح العدد السادس استخدام خوارزمية اللعبة التفاضلية ذات المحصلة الصفيرية لمعالجة مشكلة التحكم في مناورة التشكيل المتقارب لنظام متعدد الوكلاء من الدرجة الثانية يتعرض لهجمات المشغلات. في هذا المخطط، يتم التعامل مع سياسة التحكم الأمثل غير الموثوق بها وإشارات هجوم المشغل كفريقين متعارضين. يتم اشتقاق سياسات التحكم المثلى الموزعة بين القائد والتابع وسياسات التوهين للهجوم السبراني من حل توازن ناش لمعادلة هاملتون-جاكوبي-إسحاق (IJH). تم تقدير حل معادلة IJH بمساعدة الشبكة العصبية الناقدة. يتم تعلم أوزان النظام الناقد عبر الإنترنت بناءً على مخطط البرمجة الديناميكية التكيفية (PDA).

CHAPTER 1

INTRODUCTION

This chapter provides the main introduction of the thesis. The introduction consists of the background and motivation of the work, the important contributions, and the outline of the thesis.

1.1 Background

1.1.1 Multi-agent cyber-physical systems

Cyber-physical systems are multi-layer systems that integrate computations and physical components such as actuators, controllers, and sensors with communication networks [1]. Recently, cyber-physical systems have witnessed a remarkable growth in research interest as a result of their broad applications in several fields including electrical power generation and distribution, future smart grids, public health systems, transportation networks, water/gas distribution networks, and manufacturing plants [2].

Multi-agent systems as a vital category of cyber-physical systems have received

immense attention in the past years for the coordination of agents interacting together to jointly execute complex tasks that are difficult for a single agent to accomplish [3]. The agents could be of any kind, including drones, satellite clusters, robotic arms, sensor networks, vehicles, and so on. In multi-agent cyber-physical systems, each agent is provided with a sensor, local control unit, and communication devices to share information with neighboring agents over a wireless network to collaborate in a distributed manner and realize a common goal without the need for a centralized command. As a result, the following benefits are achieved [4], [5]:

1. Low cost of installation: The control functions are localized with remote monitoring and supervision thereby minimizing cost and improving reliability.
2. Scalability: The system can be easily expanded just by connecting more agents, control units, and I/O modules
3. System redundancy: In case of failure of an element in the distributed control system, the redundancy ensures that the system's operation continues and only the affected element is interrupted.
4. System security: The Distributed Control framework provides a secured system that enables effective control performance at various levels of the cyber-physical system.

From the physical point of view, multi-agent systems are broadly classified into homogeneous multi-agent systems and heterogeneous multi-agent systems [6], [7]. Agents having identical dynamics or physical forms are referred to as homogeneous. On the

other hand, agents having different dynamics or physical forms are called heterogeneous.

1.1.2 Multi-agent formation control:

Multi-agent formation control is inspired by natural behaviors such as a pack of wolves, bee colonies, bird migration, colonies of ants, and a school of fish where various formations are formed to enhance survivability, adaptation to environmental changes, efficiency in executing tasks, and defense against predators [8], [9]. Formation control is a coordinated control of a collaborative group of mobile agents (i.e. satellites, autonomous vehicles) to follow specified trajectories while maintaining the prescribed geometric shapes and distances between neighboring agents. Furthermore, formation control assists in preventing inter-agent collision and collision with obstacles [10].

Over the years, numerous formation control problems for multi-agent systems have been investigated. These problems are broadly classified into formation shape control problems and formation maneuver control problems [11]. formation shape control is the fundamental formation control problem and aims to drive networked agents to configure into a prescribed geometric shape [12]. Meanwhile, formation maneuver control improves the flexibility of the coordinated behavior of the agents. The achieved formation of the agents is set in motion in space according to a specified trajectory [13].

Moreover, formation control problems can be classified according to the method used to attain the predefined formation pattern as bearing-based, distance-based, displacement-based, and position-based.

- Bearing-based formation [14], [15]: In this approach, the mobile agents are required to maintain predefined relative bearings between them but the inter-agent distances are flexible. The flexibility in the distances enables the scaling of the formation. The agents that converge to the reference formation through the bearing-based technique can have scaling and translation formation maneuvers.
- position-based formation [16]: The agents determine their absolute positions with respect to a global coordinate system. The reference formation is described by the reference positions of the agents. The agents control their respective positions to converge to the reference formation. The resulting formation can only realize time-varying formation translation maneuvers.
- Distance-based formation [17], [18]: Here, each agent detects the relative positions of its neighbors in reference to its local coordinate system. The local coordinate systems of the agents don't have to be in a common orientation. The reference formation is defined in terms of the inter-agent distances. The reference formation is obtained by controlling the reference distances among the neighboring agents. Formations formed from the distance-based scheme can maneuver in rotation and translation.
- Displacement-based formation [16], [19]: The agents detect the relative positions of their neighbors in reference to a global coordinate system. However, the agents are not required to know their absolute position with respect to the global coordinate system. The agents actively control the displacements and form the target formation from the target displacements in reference to a global coordi-

nate system. Formations formed from the bearing-based scheme can maneuver in scaling and translation.

It is worth noting that the aforementioned formation shape control strategies cannot be used to obtain rotation, scaling, and translation maneuvers simultaneously. In addition, the strategies cannot achieve formation collinearity and shearing maneuvers. To tackle these limitations, affine formation control methods have been deployed in recent years [20], [21].

1.1.3 Multi-agent affine Formation Control

The affine formation control technique uses the properties of the stress matrix and the principle of affine transformation to track any time-varying desired formation that corresponds to an affine transformation of a constant nominal configuration [20], [21]. The stress matrix can be regarded as a generalized form of the Laplacian matrix and the weight of its edges can be negative, positive, or zero. Affine transformation is a geometric transformation that includes collinearity, rotation, scaling, shear, and translation. The benefit of the stress matrix is that it is invariant to any affine transformation applied to a nominal formation. This makes it an essential tool for affine formation control. The affine formation control provides more degree of maneuverability to a group of mobile agents.

1.1.4 Security challenge of multi-agent cyber-physical systems

The advances in technology make it possible for the cooperative operations of multi-agent cyber-physical systems with the aid of networked control systems. However, communication networks are exposed to security risks from malignant cyber-attacks [22]. These attacks may compromise data transmission among the actuators, controllers, and sensors. Consequently, the attacks can cause physical damage, loss of stability, and loss of production [23]. In essence, the malicious attacks on networked systems are of three types: deception attacks also called false data injection attacks, denial of service attacks, and replay attacks.

- Deception attacks [24], [25]: This is a type of cyber-attack in which data integrity is compromised; that is, adversaries trick network nodes into acquiring corrupted data. This type of attack can be tackled using a technique similar to the fault-tolerant approach, in which the fake data is treated as a fault signal.
- Denial of service attacks [26], [27]: This attack disrupts data transmission by feeding a large volume of unwanted signals to communication channels to prevent agents from communicating with their neighbors, which may lead to the failure of the control system due to the absence of real-time signals.
- Replay attacks [28], [29]: This attack is carried out by fraudulently delaying or repeating valid sensor data transmission over a communication network. The adversarial attackers do not need any knowledge of the developed control policy

or predictor to initiate it, and it can be addressed using counters or time-stamps in the transmitted messages.

Thus, it is imperative to ensure the security and resilience of the multi-agent cyber-physical systems. Recently, artificial intelligence-based techniques have been employed to attenuate the negative effects of harmful cyber-attacks [30].

1.1.5 Artificial intelligence in multi-agent cyber-physical systems

Artificial intelligence also referred to as machine intelligence, is a set of computer programs that simulate human cognitive abilities. Artificial intelligence algorithms are generally built to perform communication, learning, manipulating machinery, moving objects, perception, planning, processing, and reasoning [31]. Multi-agent cyber-physical systems are feasible; thus, the integration of artificial intelligence and multi-agent cyber-physical systems is also feasible [32]. This has the potential to improve the resilience and efficiency of cyber-physical systems. Artificial intelligence algorithms can detect malicious cyber threats and take appropriate action to neutralize the threats. Various control, detection, estimation, and mitigation approaches based on artificial intelligence have been proposed in the literature to protect the integrity of cyber-physical systems [32], [33], [34], [35].

1.2 Motivation

In recent years, we have witnessed the growth of autonomous systems that use embedded computers and communication networks to make autonomous decisions [36]. The technological advancement of autonomous systems has paved the way for various emerging autonomous technologies. One of those technologies, the distributed affine formation of cooperative multi-agent cyber-physical systems is the focus of this research. This technology provides efficient, fast, reliable, and versatile cooperation of multi-agent cyber-physical systems. Nonetheless, in practical applications, completely distributed multi-agent affine formations are prone to issues associated with communication networks and physical elements.

Modern multi-agent cyber-physical systems are built with Bluetooth, smart actuators, smart sensors, Wi-Fi, and wireless gateways allowing them to interact with each other, and the external world. However, the interactions are susceptible to challenges such as limited communication bandwidth, switching topologies, communication delay, and malicious cyber-attacks.

Moreover, from the physical point of view, the multi-agent systems are exposed to challenges such as environmental disturbances and uncertainties, heterogeneous dynamics arising from agents possessing different state orders and modes of motion, lack of model information in real systems, faults, and unstructured nonlinearities (such as deadzone, backlash, hysteresis, and saturation), etc.

Multiple challenges from the cyber and physical layers endanger the efficiency, resilience, robustness, and stability of the multi-agent affine formations. Furthermore,

it is difficult to have prior information about the type or magnitude of the challenges threatening the integrity of the multi-agent affine formations. Thus, designing a distributed affine formation control protocol for multi-agent cyber-physical systems to handle adverse circumstances, and maintain the reliability, resilience, and safety of the entire system is significant. Moreover, incorporating learning algorithms into the control protocol will make it model-free. This implies that the controller can detect, estimate, and counter any adversarial threat to the system without any prior knowledge about them.

1.3 Objectives of the Thesis

The aim of this research is to provide an effective distributed leader-follower affine formation maneuver control strategy for collaborative multi-agent cyber-physical systems to make significant contributions compared to existing techniques. Specifically, the core objectives of the research are outlined as follows:

1. To study the fault-tolerant control issues for leader-follower affine formation maneuver of faulty multi-agent systems interacting over an undirected graph topology.
2. To investigate the problems of leader-follower multi-agent systems subjected to malicious cyber-attacks in the midst of affine formation maneuvering.
3. Integration of artificial intelligence-based learning mechanisms into the affine formation control protocol to detect, estimate, and counter any unexpected

adverse condition.

1.4 Thesis Structure and Contributions

This section presents a description of the content of each chapter. Furthermore, this section provides the journal and conference papers published or unpublished that are used in the thesis.

- Chapter 2 introduces the notations used in the thesis, together with a brief description of graph theory and affine formation.
- Chapter 3 presents the existing literature on the recently reported affine formation maneuver control methods.
- Chapter 4 presents a reinforcement learning-based control scheme for second-order continuous-time multi-agent systems subjected to external disturbances and actuator cyber-attacks during AFMs. A long-term performance index is designed to aid the realization of the distributed leader-follower affine formation control protocol. A critic neural network is used to obtain the approximate solution of the long-term performance index which is difficult to obtain, especially for nonlinear systems subjected to cyber-attacks. An actor neural network is used to obtain the distributed control policy capable of achieving the leader-follower affine formation maneuvers and mitigating cyber-attacks. The Nussbaum functions are applied to address this issue since attacks can result in a loss of control direction. The stability of the closed-loop system has been emphasized with

the use of a Lyapunov candidate function. The performance of the suggested strategy is then supported by a numerical simulation.

- Chapter 5 presents the leader-follower affine formation maneuver (AFM) problem for uncertain nonlinear multi-agent systems, in which some complex and challenging issues, such as unknown external disturbances, unknown input backlash-like hysteresis, and sensor deception attacks are addressed with the aid of the actor-critic learning control based on neural networks. The main control objective is to steer the group of followers to track the maneuvers of the group of leaders within a finite time. The control protocol is built using a distributed fast terminal sliding mode surface equipped with a reaching law to facilitate finite-time convergence. The actor neural network is handed the responsibility of estimating the uncertain nonlinear dynamics, unknown input backlash-like hysteresis, and deception attack signals and compensating them through the control protocol. Meanwhile, the critic neural network evaluates the operation of the control protocol. Lyapunov stability theorem is deployed to prove that the closed-loop system is bounded in the vicinity of zero in finite time. Simulation results are provided to demonstrate the effectiveness of the proposed control scheme.
- Chapter 6 studies the design of a novel prescribed-performance controller for fault-tolerant secure optimized backstepping control using reinforcement learning for affine formation maneuvering of second-order nonlinear multi-agent systems under sensor deception attacks and actuator faults. The control objective

is to ensure that the leader-follower maneuvers are resilient in the event of deception attacks and actuator faults. Linear transformation matrices are used to obtain various formation maneuvers such as translation, rotation, scaling, shearing, and collinearity simultaneously. Whenever the desired formation maneuver changes, the dynamic motion of the agents undergoes new transient states before converging to the new formation maneuver. In order to constrain the transient states of the formation tracking errors to a predefined small set whenever desired formation maneuver changes, a modified prescribed performance function with finite settling time is introduced. The optimal backstepping control laws are obtained from the approximate solution of the Hamilton-Jacobi-Bellman equation using reinforcement learning under the identifier-actor-critic neural networks. The resulting control algorithm is simple and does not require complex computations. Lyapunov candidate functions have been used to prove the boundedness of the closed-loop system. Simulation results have shown that the proposed controller can achieve various leader-follower affine formation maneuvers regardless of the sensor deception attacks and actuator faults.

- In chapter 7, an actor-critic learning scheme for safe leader-follower affine formation maneuver control of networked quadrotors is studied, in the face of external disturbances, sensor deception attacks, and injection attacks on the actuators. The followers aim to track the formation maneuvers such as scaling, shearing, translation, and rotation decided by the leaders. Motivated by increasing safety and performance requirements during formation maneuvering,

the dynamic states of the quadrotors are constrained within the prescribed safety region. To guarantee that the safety constraints are not violated, a barrier Lyapunov function is put in place. Then, a distributed sliding mode control with actor-critic learning is formulated to facilitate accurate leader-follower affine formation maneuvers and reject malicious cyber-attack signals. The input gains that arose due to the attacks might corrupt the control direction. The Nussbaum gain function is coupled to the controller to tackle this problem. The actor system is responsible for estimating the uncertain dynamics together with the malicious attack signals while the critic network evaluates the control performance through the estimated long-term performance index. The overall stability of the closed-loop system has been proven to be uniformly ultimately bounded using the Lyapunov stability theorem. Finally, simulation results demonstrate the capability of the presented control method.

- In Chapter 8, a distributed control law with a neural network is designed to achieve leader-follower affine formation maneuver (AFM) of multi-agent systems (MASs) with adversarial disturbances. The leaders determine the desired collective formation maneuvering such as scaling, shearing, translation, and rotation. The distributed controller ensures that the followers are tracking the maneuvers of the leaders at all times despite the disturbances. The disturbances are approximated and compensated with the aid of neural networks. A Lyapunov candidate function is used to highlight that the closed-loop system is uniformly ultimately bounded. Finally, in order to validate the proposed control method,

it is applied to a multi-agent system of quadrotors. The simulation results have shown that the proposed control protocol is able to maintain the collective affine formation maneuvers.

- In chapter 9 the affine formation maneuver (AFM) control problem using the zero-sum differential game is studied for a second-order multi-agent system (MAS) against actuator cyber-attacks. First and foremost, the distributed control signals and the cyber-injected attack signals are viewed as two competing teams. Then, the leader-follower tracking error system is established where the followers aim to track the AFMs of the leaders. The leader-follower distributed optimal control policies and the cyber-attack attenuation policies are derived from the Nash equilibrium solution of the Hamilton-Jacobi-Isaac (HJI) equation. The solution of the HJI equation is estimated with the aid of the critic neural network. The weights of the critic system are learned online based on the adaptive dynamic programming (ADP) scheme. The uniformly ultimately bounded stability of the closed-loop system under the formulated zero-sum differential game control strategy has been proven by the Lyapunov stability theorem. A simulation example illustrates that the presented control method can attenuate actuator cyber-attacks and maintain the AFMs of the agents.

1.5 Publications

- S. El-Ferik, M. Maaruf, F. M. Al-Sunni, A. A. Saif and M. M. Al Dhaifallah, "Reinforcement Learning-Based Control Strategy for Multi-Agent Systems Sub-

- jected to Actuator Cyberattacks During Affine Formation Maneuvers,” in IEEE Access, vol. 11, pp. 77656-77668, 2023, doi: 10.1109/ACCESS.2023.3296741.
- M. Maaruf, S. El-Ferik and F. M. Al-Sunni, ”Neural Network-Based Control for Affine Formation Maneuver of Multi-Agent Systems with External Disturbances,” 2023 31st Mediterranean Conference on Control and Automation (MED), Limassol, Cyprus, 2023, pp. 862-867, doi: 10.1109/MED59994.2023.10185905.
 - M. Maaruf, S. El-Ferik, A. A. Saif, F. M. Al-Sunni and M. M. Al-Dhaifallah ”Affine Formation Maneuvering of Nonlinear Multi-Agent Systems with Fault-Tolerant Secure Optimized Backstepping Control Using Reinforcement Learning”, (IEEE Transaction on Neural Networks and Learning Systems, under review).
 - M. Maaruf and S. El-Ferik, ”Learning-Based Safety-Guaranteed Sliding Mode Affine Formation Maneuver Control of Quadrotors Vulnerable to Cyber-Attacks”, (Aerospace Science and Technology, under review).
 - S. El-Ferik and M. Maaruf, ”Critic Learning Control Via Zero-Sum Differential Game for Affine Formation Maneuver of Multi-agent Systems with Cyber-Attacks”, (Expert Systems, under review).
 - M. Maaruf and S. El-Ferik, ”Actor-critic Learning-Based Finite-Time Secure Formation Maneuver Control of Networked Systems with Deception Attacks and Backlash-like Hysteresis”, (To be submitted to Neurocomputing).

CHAPTER 2

PRELIMINARIES

2.1 Notations

This thesis utilizes the standard notations commonly used in the literature. For instance, $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n \times m}$ represent real numbers, n – dimensional real vector, and $(n \times m)$ – dimensional real matrix, respectively. We use I_n and $\mathbf{1}_N$ to denote $(n \times n)$ – dimensional identity matrix and N – dimensional vector of ones, respectively. The symbol $\|\cdot\|$ is used to denote the Frobenius norm. The symbol $\otimes$ denotes the Kronecker product.

2.2 Graph Theory

The interaction among the agents in the cyber layer is modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ with N nodes, where $\mathcal{V} = \{v_1, \dots, v_N\}$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, $\mathbf{A} = [\mathbf{a}_{ij}]$ is the adjacency matrix with weights $\mathbf{a}_{ij} = 1$ if $(\mathbf{v}_i, \mathbf{v}_j) \in E$ and $\mathbf{a}_{ij} = 0$ if $(\mathbf{v}_i, \mathbf{v}_j) \notin E$. The laplacian matrix $\mathbf{L} = [\mathbf{l}_{ij}]$ of the graph $\mathbf{G}$ is given

by $\mathbf{l}_{ij} = -\mathbf{a}_{ij}$, $i \neq j$ and $\mathbf{l}_{ij} = \sum_{j=1, j \neq i}^N \mathbf{a}_{ij}$, $i \neq j$. The set of neighbors of node i is represented by $\mathcal{N}_i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ [37].

2.3 Formation Graph

The configuration of the N agents in n – dimensional Euclidean space is define by $q = [q_1^T \ q_2^T \ \dots \ q_N^T]^T \in \mathbb{R}^{nN}$ with $q_i \in \mathbb{R}^n$. The formation of the agents, denoted by $(\mathcal{G}, q)$, is the interaction graph along with the corresponding configuration of the agents [16].

Considering that this thesis is interested in leader-follower affine formation control, the set of vertices of the formation $\mathcal{V}$ can be partitioned into leaders' subset $\mathcal{V}_l = \{1, \dots, N_l\}$, and followers' subset $\mathcal{V}_f = \mathcal{V} \setminus \mathcal{V}_l = \{N_l + 1, N_l + 2 \dots, N\}$. Consequently, the configuration of agents can be partitioned as $q = [q_l^T \ q_f^T]^T$, where

$$q_l = [q_1^T \ q_2^T \ \dots \ q_{N_l}^T]^T \in \mathbb{R}^{nN_l} \quad (2.1)$$

$$q_f = [q_{N_l+1}^T \ q_{N_l+2}^T \ \dots \ q_{N_l+N_f}^T]^T \in \mathbb{R}^{nN_f} \quad (2.2)$$

2.4 Affine Formation Transformation

The target planar shape that the agents are required to form and keep is represented by the constant configuration $v = [v_1^T \ v_2^T \ \dots \ v_N^T]^T = [v_l^T \ v_f^T]^T \in \mathbb{R}^{nN}$ under a nominal formation $(\mathcal{G}, v)$. The nominal formation $(\mathcal{G}, v)$ can be affinely transformed into a new time-varying formation $(\mathcal{G}, q(t))$ with the aid of linear transformation matrices. The affine transformation of v is calculated as follows [21]:

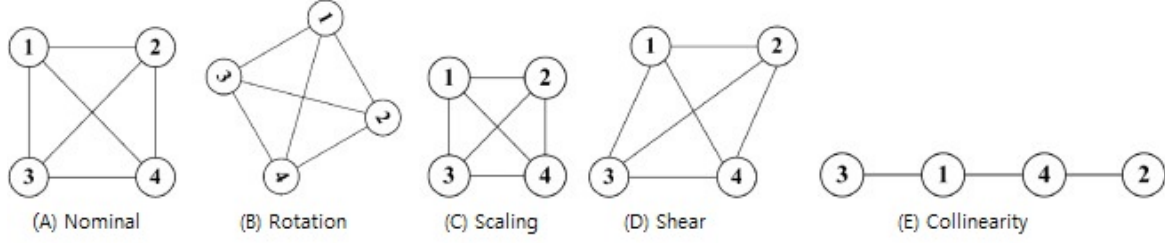


Figure 2.1: Various affine transformations of a nominal configuration.

$$q(t) = [I_N \otimes A(t)]v + 1_N \otimes b(t) \quad (2.3)$$

Similarly, for each agent, the affine transformation of $v_i \in \mathbb{R}^n$ is thus:

$$q_i(t) = A(t)v_i + b(t) \quad (2.4)$$

where $b(t) \in \mathbb{R}^n$ stands for the translation maneuver of the formation, $A(t) \in \mathbb{R}^{n \times d}$ can be manipulated to achieve the rotation, scaling, and shearing maneuvers of the whole formation with respect to v . Fig. 2.1 illustrates several formation shape maneuvers of the agents.

2.4.1 Stress matrix

The stress matrix of the formation $(\mathcal{G}, q)$ is a set of scalar weights $\{\omega_{ij}\}_{(i,j) \in \mathcal{E}}$ assigned to each edge such that $\omega_{ij} = \omega_{ji} \in \mathbb{R}$ [8]. The stress matrix $\Omega \in \mathbb{R}^{N \times N}$ can be defined as [21]:

$$[\Omega]_{ij} = \begin{cases} 0, & i \neq j, (i, j) \notin \mathcal{E} \\ -\omega_{ij}, & i \neq j, (i, j) \in \mathcal{E} \\ \sum_{k \in \mathcal{N}_i} \omega_{ik}, & i = j. \end{cases}$$

A stress matrix is said to be equilibrium stress when:

$$\sum_{j \in \mathcal{N}_i} \omega_{ij}(q_j - q_i) = 0, \quad i \in \mathcal{V} \quad (2.5)$$

Equation (2.5) can be transformed to a matrix as:

$$(\Omega \otimes I_n)q = 0 \quad (2.6)$$

The stress matrix Ω can be partitioned as

$$\bar{\Omega} = \Omega \otimes I_n = \begin{bmatrix} \Omega_{ll} & \Omega_{lf} \\ \Omega_{fl} & \Omega_{ff} \end{bmatrix} \otimes I_n = \begin{bmatrix} \bar{\Omega}_{ll} & \bar{\Omega}_{lf} \\ \bar{\Omega}_{fl} & \bar{\Omega}_{ff} \end{bmatrix}. \quad (2.7)$$

where $\Omega_{ll} \in \mathbb{R}^{nN_l \times nN_l}$, $\Omega_{ff} \in \mathbb{R}^{nN_f \times nN_f}$ and $\Omega_{fl} \in \mathbb{R}^{nN_l \times nN_f}$

Definition 2.1 [21] **Affine span** : For a given set of points $\{q_i\}_{i=1}^N \in \mathbb{R}^n$, the affine span (S) of these points is defines as:

$$\mathcal{S} = \left\{ \sum_{i=1}^N a_i q_i : a_i \in \mathbb{R} \text{ for all } i \text{ and } \sum_{i=1}^N a_i = 1 \right\}. \quad (2.8)$$

The affine transformations of v are contained in the image of v defined as follows:

$$\mathcal{A}(v) = \{q \in \mathbb{R}^{nN} : q = (I_N \otimes A)v + 1_N \otimes b\}. \quad (2.9)$$

Definition 2.2 [38] **Affine localizability** : If the target position of the followers q_f can be uniquely calculated from the position of the leaders q_l for any $q = [q_f^T \ q_l^T]^T \in \mathcal{A}(v)$, then the nominal formation $(\mathcal{G}, v)$ is affinely localizable by q_l .

Assumption 2.1 [39] *For the design of the formation maneuver control, the following assumptions are essential.*

1. *The formation graph $(\mathcal{G}, v)$ is universally rigid.*
2. *For motion in an n -dimensional space, the number of leaders must be selected such that $N_l = n + 1$ to affinely span the entire n -dimensional space.*
3. *The nominal configuration v is affinely spanned in $\mathbb{R}^n$.*
4. *The equilibrium stress matrix Ω is semi-definite such that $\text{rank}(\Omega) = N - n - 1$.*
5. *The position q_f is affinely localizable by q_l .*

Considering Definition 2.2 and Assumption 2.1, for any $q = [q_f^T \ q_l^T]^T \in \mathcal{A}(v)$, q_f can be uniquely calculated from q_l as follows:

$$q_f = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_n) q_l. \quad (2.10)$$

CHAPTER 3

LITERATURE REVIEW

This chapter gives a review of the exiting affine formation maneuver control methods reported in the literature.

3.1 Existing Literature on affine formation maneuver control

The first result of affine formation control using the properties of the stress matrices and the principles of affine transformation was reported in [20]. The authors discovered sufficient and necessary conditions to steer the agents to acquire the affine formation over directed and undirected communication graphs. Then, Shiyu [21] modified the technique in [20] and developed a new and generalized approach that not only steers the agents to the target formation but also maneuvers the whole formation. In this scheme, the agents are divided into two groups; the leaders and the followers. The affine transformation of the followers is uniquely obtained from the coordinates of the

leaders. Here, only the leaders have access to the target formation. A distributed control policy makes the followers track the maneuvers of the leaders and ensures the global stability of the whole formation.

In the past years, the work in [21] has inspired several publications on affine formation maneuver control. These are summarized as follows:

3.1.1 Affine formation maneuver control of MASs described by integrator dynamic equations

The authors in [38], [39], designed an affine formation algorithm for linear systems governed by triple integrator dynamic equations in both continuous-time and sampled-data modes. To improve the convergence of the followers to the affine positions localized by the leaders, a prescribed convergence time function is integrated into the distributed controller for single-integrator agents [40], [41], [42]. In [43], the problem affine formation control of single-integrator systems considering the optimal topology design has been studied. In [44], an affine formation maneuver control of double-integrator MASs with external disturbances. In [45], a linear matrix inequalities-based control law is implemented for a MAS described by double-integrator models. In [46], the affine formation maneuver control problem of double-integrator MASs with prescribed convergence time has been studied. the authors studied the potential of affine formation maneuver control for application in inspecting transmission lines. In [47], An adaptive component is added to the controller to estimate and suppress external disturbances. In [48], an improved affine formation maneuver controller for

second-order integrator MASs is designed using the acceleration history. In [49], a proportional-integral (PI) controller is used to achieve the affine formation maneuver of high-order-integrator MASs with time-varying communication delays. In [50], the authors investigated the affine formation maneuver of high-order MASs using a two-layered approach. In this approach, the agents are divided into three categories; the first leader, the second leader category, and the followers. The first leader will determine the type of FM while the second leader group will retain the desired positions relative to the first leader using distributed adaptive finite-time control. Then, a distributed proportional-derivative (PD) controller with adaptive gains is used to move the followers to the target formation determined by the first leader. Motivated by [50], Fang et al. [51] investigated the FM control of chain integrator MASs connected by a directed graph. Here, proportional controllers and proportional-derivative controllers were developed for the leaders to guarantee global convergence of the leaders to the desired trajectories. Similarly, proportional controllers and proportional derivative controllers were designed for followers to guarantee global convergence of the followers to the desired maneuvers of the leaders. In [52], an improved affine formation maneuver control was designed for second-order integrator systems considering the acceleration history. The parameters of the control protocol are derived on the basis of the neutral differential equations. In [53], an enhanced affine formation maneuver control of multi-vehicle systems was presented by using a historical velocity command. The authors presented experimental results to support the theoretical results. The authors in [54] studied the stability analysis of a second-order leader-follower multi-agent system with multiple time delays. In [55], a distributed Kalman filter was used

to estimate the states of single-integrator agents which were used to design the affine formation controller. In [56], a distributed output feedback affine formation maneuver control of leader-follower single-integrator multi-agent systems has been presented. [57], a new leader-follower distributed control technique was proposed for agents described with single-integrator dynamics to realize time-varying maneuvers. In [58], the affine formation maneuver control of chain integrator systems in both two-dimensional and three-dimensional spaces is studied.

3.1.2 Affine formation maneuver control of MASs described by general linear dynamic equations

A network of integrators is a special form of a linear system whose coefficients are unity and cannot represent physical systems. An alternative solution to this issue is to use a general linear system whose coefficients are not unity and can be modified to represent any physical system. In [59], the authors investigated the distributed affine formation maneuver control problem of general linear MASs with dynamic coupling gains and dynamic uncertainties. Then, robust adaptive control protocols have been presented to take care of the gains and uncertainties of the dynamic coupling. In [60], [61], an optimal control strategy has been presented for affine formation maneuver control of general linear multi-agent systems. In [62], a linear quadratic regulator strategy is used to guide followers to acquire the desired geometric shape and the maneuvers of the leaders. The proposed method in [63] utilized an output feedback control to steer networked high-order chain integrator dynamic agents to achieve several formation

shapes and maneuvers. In [64], event-triggered affine formation maneuver control of linear MASs. In [65], an inverse optimal control protocol is designed to minimize the fuel consumption of the followers during convergence to the FM target localized by the leaders. In [66], the followers with discrete-time general linear dynamics are communicating over a discrete-time directed graph. Then, an event-triggered mechanism is added to the affine formation maneuver controller whose gains are estimated online. In [67], a system of heterogeneous linear systems is controlled such that various leader-follower affine formations are obtained in the presence of external disturbances.

3.1.3 Affine formation maneuver control of MASs described by nonlinear dynamic equations

In [68], the authors investigated the affine formation maneuver control of uncertain second-order nonlinear MASs. The nonlinearities of the agents were parametrized and adaptive laws were used to estimate the uncertain parameters. Moreover, proportional-integral controllers are employed to guarantee that the followers attain the required formation of the leaders. In addition, the gains of the controllers are adjusted online for smooth maneuvers. In [69], a sliding manifold-like variable is introduced to formulate a distributed affine formation maneuver control of marine surface vehicles. To guarantee the localization of the cooperation, a time-varying constraint is designed to monitor the attitude and position simultaneously. In [70], the affine formation controller is integrated with estimators to estimate the target maneuvers of the leaders. In [71], a distributed sliding mode control approach is used to realize

the affine transformation maneuver of collaborative autonomous surface vehicles. The authors incorporated an event-triggered data transmission mechanism to lessen the communication period among the agents. In [72], an adaptive local tracking controller is implemented for affine transformation maneuver of collaborative underactuated surface vessels. The complex nonlinear dynamics of the vessels together with the external disturbances are estimated with the aid of neural networks. Then, a distributed event-triggered observer is used to estimate the local reference state of each follower-vessel instead of the global states of the leaders. In [73], an adaptive distributed backstepping control triggered by events of underactuated surface vessels subjected to external disturbances has been presented. In [74], the adaptive affine formation control of satellites under parametric uncertainties and external disturbances has been presented. In [75], robust distributed control of heterogeneous autonomous surface vehicles with external perturbations is studied for transformations of the affine formation of the leader-follower. In [76], the affine transformation maneuver of unmanned surface vessels is studied considering the rigidity theorem. In [77], three-dimensional control for affine formation of underactuated quadrotors has been investigated. The authors in [78], [79] studied the layered affine formation control of nonlinear systems to achieve several formation transformations. In [80], the affine formation control of unmanned aerial vehicles with asymmetrical velocity constraints is studied. In [81], prescribed-time affine formation control of unmanned aerial vehicles is introduced. In [82], an affine formation control of a system of underactuated nonholonomic agents is introduced.

3.2 Research gaps

So far, the affine formation maneuver control of multi-agent systems has not been fully explored. All in all, there are a lot of control problems that have not yet been investigated with respect to affine formation maneuvers, such as:

1. The problem of malicious cyber-attacks on the agents during formation maneuvering has not been studied.
2. Even though, MASs are susceptible to faults, no literature has reported the fault-tolerant control of MASs in affine-transformation maneuver.
3. Artificial intelligence-based techniques (such as deep learning, reinforcement learning, and data-driven approaches) are yet to be deployed for application in affine formation maneuvers of MASs.
4. Prescribe performance control has been widely applied to various dynamic systems, but is yet to be used to achieve affine formation maneuvering with preassigned performance.

3.3 Proposed control problem

Considering the areas that have not been investigated with respect to the affine formation maneuver of MASs based on stress matrix and the principles of affine transformations, this research proposes learning-based control schemes using artificial intelligence to achieve affine formation maneuvers of leader-follower MASs subjected to cyber-attacks and faults. The learning-based algorithms enable the controllers to estimate the magnitude of the cyber-attacks and faults and then take appropriate action to mitigate them. Moreover, the controllers become model-free as they require no information about the systems' dynamics.

To the best of our knowledge, the proposed control problem has not been addressed in the literature.

CHAPTER 4

REINFORCEMENT LEARNING-BASED CONTROL STRATEGY FOR MULTI-AGENT SYSTEMS SUBJECTED TO ACTUATOR CYBERATTACKS DURING AFFINE FORMATION MANEUVERS

This chapter provides the reinforcement learning-based control strategy for second-order continuous-time multi-agent systems (MASs) subjected to actuator cyberattacks during affine formation maneuvers. In this case, a long-term performance index is created to track the MASs tracking faults using a leader-follower structure. In order to approximate the ideal solution, which is challenging to find for systems vulnerable to cyberattacks during time-varying maneuvers, a critical neural network is used. The distributed control protocol is obtained, and the long-term performance index is minimized, using an actor neural network strengthened with critic signals. The actor-critic neural networks calculate unknown dynamics and the severity of attacks on the MAS actuators. The Nussbaum functions are applied to address this issue since attacks can result in a loss of control direction. The stability of the closed-loop system has been emphasized with the use of a Lyapunov candidate function. The performance of the suggested strategy is then supported by a numerical simulation.

4.1 Introduction

MASs are increasingly connected to wireless networks where information transmission is facilitated for remote supervision and control. As such, they are prone to malicious cyber attacks such as deception attacks, denial of service attacks, false data injection, integrity attacks, malware attacks, replay attacks, and so on [83]. These attacks mostly inject misleading signals into the systems' sensors, actuators, or communication channels, which may lead to the deterioration of the performance of the systems [84]. There exist several control strategies for detecting and mitigating various cyber attacks to secure the consensus of MASs [85], [86], [87], [88], [89] or to secure the geometrical formation acquired by the agents [90], [91], [92], [93], [94], [95], [96], [97]. However, a secure control against cyber attacks is yet to be investigated with respect to AFM of MAS. A false data injection attack on the actuators of the leaders or the followers will lead to loss of collective maneuver and the agents may collide. Therefore, how to secure the leaders as well as the followers so that collective formation maneuvering is maintained in adversarial environments is an interesting area to explore.

As an intelligent technique, a reinforcement learning (RL) scheme has been widely employed for solving optimal adaptive control problems by minimizing a prescribed cost function to maximize a reward received from the environment [98]. For continuous-time systems, integral RL is utilized to compute the solution of the optimal adaptive control problem using the actor-critic framework. Considering the fact that it is difficult to obtain the solution of the optimal adaptive control problem of continuous-time systems using RL, neural networks are commonly used to compute

the approximate solution. Approximate RL-based control has been applied to several classes of single agent systems [99], [100], [101], [102], [103], MAS[104], [105], [106], [107]. Recently, approximate RL is employed for formation acquisition control of MASs [108], [109], [110], [111], [112]. The problem of AFM control of MASs using approximate RL deserves to be investigated and related work is not reported yet.

In the existing RL control articles, the AFM control problem of MASs has not been studied yet, let alone simultaneous consideration of malicious cyber attack cases. Motivated by this observation, this article proposes an RL scheme using the actor-critic framework to achieve the AFM control of general nonlinear MASs in an adversarial environment. Due to the difficulties in obtaining the optimal solution of the long-term performance function for the AFM of MASs with actuator cyber attacks, critic neural network is used to obtain the approximate solution. The critic evaluates the control performance and sends the reinforcement signals to the actor. Then, the actor applies the control signals to the actuators of the agents based on the critic evaluation. Here, two actor networks are used: one for estimating the uncertain nonlinear dynamics of the agents, and the other for estimating the actuator attacks. The attacks on the actuator may alter the control direction of the distributed control protocols. This problem is addressed by using the Nussbaum functions to estimate the control direction. To the authors' best knowledge, the proposed control problem has not been investigated by anyone yet. This chapter differs from existing AFM control methods as follows:

1. Unlike the leader-follower AFM control approaches in [68], [21], an RL control

based on actor-critic neural networks is proposed to achieve the AFM of the leader-follower MAS in adversarial environments.

2. The existing AFM control protocols will fail to maintain the leader-follower AFM when subjected to cyber attacks on the actuators. Here, this problem is tackled by using the actor neural network to learn and counter the actuator attacks.
3. In [68], the authors considered a class of second-order uncertain nonlinear MAS that can be expressed in the linear-in-the-parameter form, and adaptive laws were used to estimate the uncertain parameters. However, in this study, the uncertain nonlinear MAS doesn't need to be linear-in-the-parameter form since an action neural network is used to approximate it. Moreover, the action neural network which receives reinforced signals from the critic neural network enhances the learning and optimization of the system.
4. The loss of control direction as a result of the actuator attack is tackled using the Nussbaum function.

4.2 Problem formulation

4.2.1 Dynamic models of the agents

Consider second-order nonlinear multi-agent systems consisting of N_l leaders and $N_f = N - N_l$ followers. The nonlinear dynamics of the N_f followers can be described as:

$$\begin{cases} \dot{q}_i = v_i \\ \dot{v}_i = f_i(q_i, v_i) + u_i^c(t) + \varphi_i(q_i, v_i), \end{cases} \quad i \in \mathcal{V}_f \quad (4.1)$$

where $q_i \in \mathbb{R}^d$ and $v_i \in \mathbb{R}^d$ denote the position and velocity of the i th agent, $u_i^c(t) \in \mathbb{R}^d$ is the compromised actuator output of agent i , $f_i(q_i, v_i) \in \mathbb{R}^d$ is a smooth nonlinear function, $\varphi_i(q_i, v_i) \in \mathbb{R}^d$, $\mathcal{V}_f = \{N_l + 1, N_l + 2, \dots, N_l + N_f\}$ is the group of followers.

The trajectories of the N_l leaders are given by:

$$q_i = A(t)r_i + b(t), \quad i \in \mathcal{V}_l \quad (4.2)$$

where $\mathcal{V}_l = \{1, 2, \dots, N_l\}$ is the group of the leaders. The position of the leaders $q_l = [q_1^T \ q_2^T \ \dots \ q_{N_l}^T]^T$ is known in advance and various maneuvers can be obtained by manipulating $A(t)$ and $b(t)$.

4.2.2 Neural network approximations

Owing to the function approximation property of the radial basis function neural network (RBFNN), it has been widely deployed to estimate unknown nonlinear functions.

The estimation of continuous nonlinear function with RBFNN is expressed as:

$$\zeta_{nn} = W^T \Psi(X) + \epsilon(X) \quad (4.3)$$

where $W = [w_1 \ w_2 \ \dots \ w_m]^T \in \mathbb{R}^m$ is the ideal with vector with m number of neurons, $\epsilon(X)$ is the function estimation error satisfying $\|\epsilon(X)\| \leq \bar{\epsilon}$, $\Psi(X) =$

$[\Psi_1(X) \ \Psi_1(X) \ \dots \ \Psi_m(X)]^T$ is the Gaussian basis function vector and $\Psi_i(X)$ is

$$\Psi_i(X) = \exp\left(-\frac{(X - c_i)(X - c_i)}{\eta_i^2}\right), \quad i = 1, 2, \dots, m \quad (4.4)$$

where c_i is the center of the receptive field, η_i is the width of the Gaussian function.

4.2.3 Actuator attacks

The malignant cyber attacks considered here are the false data injection attacks into the actuator. These attacks could corrupt the control command of the i th follower.

The actuator attacks can be modelled as a state-dependent function as follows:

$$u_i^c(t) = g_i(q_i, v_i)u_i(t) + \delta_i(t, q_i, v_i) \quad (4.5)$$

where $u_i^c(t)$ is the corrupted actuator output, $u_i(t)$ is the control signal, $\delta_i(t, q_i, v_i)$ is the time-varying state-dependent injected attack, $g_i(q_i, v_i)$ is a state-dependent unknown gain causing loss of effectiveness of the actuator as a result of the attack.

The system (4.1) under the actuator attacks (4.5) is rewritten in the following form:

$$\begin{cases} \dot{q}_i = v_i \\ \dot{v}_i = f_i(q_i, v_i) + g_i(q_i, v_i)u_i(t) + \delta_i(t, q_i, v_i) + \varphi_i(q_i, v_i) \end{cases} \quad (4.6)$$

Assumption 4.1 [9] *The unknown diagonal matrix $g_i(q_i, v_i) \in \mathbb{R}^{d \times d}$ resulting from the false data injection satisfies $\underline{g}_i < \lambda_1(g_i(q_i, v_i)) < \lambda_2(g_i(q_i, v_i)) < \infty$, with $\underline{g}_i$ being a constant lower bound of $g_i(q_i, v_i)$, and λ_1 and λ_2 are the Eigen values of $g_i(q_i, v_i)$.*

Assumption 4.2 [113] *The disturbance vector φ_i is unknown and satisfies $\|\varphi_i\| \leq \rho_i$, with ρ_i being an unknown constant.*

Remark 4.1 *The false data injection model in [114], [115], [116] is of the form $u_i^c = u_i + \delta_i$, i.e $g_i = 1$. This article extends the results to the situation where the unknown gain g_i is bounded, which can accurately highlight the severity of the attacks on the actuator.*

Remark 4.2 *Assumption 4.1 is widely used in the literature to simplify the complexity of the control law [9], [10]. Instead of formulating a complex algorithm to estimate the unknown $g_i(q_i, v_i)$, using its known upper bound in the control law simplifies the control design process. Assumption 4.2 is a standard and common assumption in the robust control literature [113], [117]. This assumption is necessary because unbounded disturbances will grow and make the state trajectories of a system explode, i.e. approach infinity.*

Definition 4.1 [118] *Any continuous function $N(\beta) : \mathbb{R} \longrightarrow \mathbb{R}$ is a Nussbaum function if it satisfies the following properties:*

$$\lim_{\theta \rightarrow \infty} \sup \frac{1}{\theta} \int_0^\theta N(\beta) d\beta = +\infty \quad (4.7)$$

$$\lim_{\theta \rightarrow \infty} \inf \frac{1}{\theta} \int_0^\theta N(\beta) d\beta = -\infty \quad (4.8)$$

For example, the functions $\beta^2 \cos(\beta)$ and $\beta^2 \sin(\beta)$ are Nussbaum functions. In this work, we choose $\beta^2 \cos(\beta)$.

Lemma 4.1 *Let $\beta(t)$ be a smooth function on $[0, t_f)$, $L(t)$ be non-negative smooth function, and $N(\beta)$ be a smooth Nussbaum function. If the following inequality is valid*

$$L(t) \leq a_0 + e^{-a_1 t} \int_0^t [g(\cdot)N(\beta) + 1] \dot{\beta} e^{a_1 \tau} d\tau \quad (4.9)$$

where $a_0 > 0$, $a_1 > 0$ are constants, $g(\cdot)$ is a time-varying function, then, $L(t)$, $\beta(t)$, and $\int_0^t g(\cdot)N(\beta)\dot{\beta}d\tau$ must be bounded on $[0, t_f]$

4.3 Control design

The target positions of the followers relative to the leaders can be expressed as:

$$q_f^*(t) = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_d) q_l. \quad (4.10)$$

The control objective of the followers under the actuator attacks is thus:

$$\lim_{t \rightarrow \infty} (q_f(t) - q_f^*(t)) = \quad (4.11)$$

$$\lim_{t \rightarrow \infty} (q_f(t) + (\Omega_{ff}^{-1} \Omega_{fl} \otimes I_d) q_l) = 0 \quad (4.12)$$

Define the following error variables for the followers:

$$\xi_i = \sum_{j=1}^N \omega_{ij} (q_i - q_j) + \Lambda \sum_{j=1}^N \omega_{ij} (v_i - v_j), \quad i \in \mathcal{V}_f \quad (4.13)$$

Let $e_i = \sum_{j=1}^N \omega_{ij} (q_i - q_j)$. Therefore, one has $\xi_i = e_i + \Lambda \dot{e}_i$, $i \in \mathcal{V}_f$. Taking the time-derivatives of ξ_i , $i \in \mathcal{V}_f$, one can gets:

$$\begin{aligned}
\dot{\xi}_f &= \dot{e}_f + \Lambda \ddot{e}_f \\
&= \dot{e}_f + \Lambda \left\{ \bar{\Omega}_{ff} (f(q, v) + g(q, v)u(t) + \delta(t, q, v) \right. \\
&\quad \left. + \varphi(q, v)) + (\Omega_{fl} \otimes I_d) \ddot{v}_l \right\}
\end{aligned} \tag{4.14}$$

where $\xi_f = [\xi_{N_l+1}^T \ \xi_{N_l+1}^T \cdots \xi_{N_l+N_f}^T]^T$, $e_f = [e_{N_l+1}^T \ e_{N_l+1}^T \cdots e_{N_l+N_f}^T]^T$, $f = [f_{N_l+1}^T \ f_{N_l+2}^T \cdots f_{N_l+N_f}^T]^T$, $u = [u_{N_l+1}^T \ u_{N_l+2}^T \cdots u_{N_l+N_f}^T]^T$, $\delta = [\delta_{N_l+1}^T \ \delta_{N_l+2}^T \cdots \delta_{N_l+N_f}^T]^T$, $\varphi = [\varphi_{N_l+1}^T \ \varphi_{N_l+2}^T \cdots \varphi_{N_l+N_f}^T]^T$.

4.3.1 Critic design

A binary utility function $p(\xi_i(t))$, which is the current system performance index is set as [103]:

$$p(\xi_i(t)) = \begin{cases} 0 & \text{if } \xi_i(t)^2 \leq th_r \\ 1 & \text{if } \xi_i(t)^2 > th_r \end{cases} \quad \forall \xi \in [t - T, t) \tag{4.15}$$

where $th_r > 0$ is a designed threshold. $p(\xi_i(t)) = 0$ indicates acceptable tracking accuracy whereas $p(\xi_i(t)) = 1$ indicates unacceptable tracking performance.

The long-term performance index is described by:

$$I(t) = \int_t^\infty \alpha^{\frac{-\pi+t}{\tau}} p(\xi(\pi)) d\pi \tag{4.16}$$

where $\alpha \in (0, 1)$ is a constant that discounts the future cost, and τ is a small integral reinforcement interval, $p(\xi) = [p(\xi_{N_l+1})^T \ p(\xi_{N_l+2})^T \cdots p(\xi_{N_l+N_f})^T]^T$. If ξ_i is within the

allowable threshold, then the control objective is realized and $I(t)$ will not increase. However, if $\xi_i(t)$ is outside the allowable threshold, then the controller should be fine-tuned to make e_f , $\dot{e}_f$, and $I(t)$ smaller.

From (4.16), the continuous time Bellman error equation can be constructed as:

$$\begin{aligned}
I(t - \tau) &= \int_{t-\tau}^{\infty} \alpha^{\frac{-\pi+t-\tau}{\tau}} p(\xi(\pi)) d\pi \\
&= \alpha^{-1} I(t) + \int_{t-\tau}^t \alpha^{\frac{-\pi+t-\tau}{\tau}} p(\xi(\pi)) d\pi \\
&\triangleq \alpha^{-1} (I(t) + p_c)
\end{aligned} \tag{4.17}$$

where $p_c = \int_{t-\tau}^t \alpha^{\frac{-\pi+t-\tau}{\tau}} p(\xi(\pi)) d\pi$ is the cost value for $[t - \tau, t)$, $p_c = [p_{c_1}^T \ p_{c_2}^T \ \dots \ p_{c_{N_l}}^T]^T$ with

$$p_{c_i} = \int_{t-\tau}^t \alpha^{\frac{-\xi+t}{\tau}} p(\xi_i(\pi)) d\pi \tag{4.18}$$

$$= \begin{cases} 0 & \text{if } \xi_i(t)^2 \leq c_{p_i} \\ \frac{\tau}{\ln(\alpha)}(\alpha - 1) & \text{if } \xi_i(t)^2 > c_{p_i} \end{cases} \tag{4.19}$$

Therefore, $0 < p_{c_i} < \tau(\alpha - 1)/\ln(\alpha)$.

Given that $I(t)$ contains the future information of the system with actuator attacks, it is difficult to obtain its solution. A critic RBFNN is introduced to approximate it.

$$I(t) = W_c^T \Psi_c(X_c(t)) + \epsilon_c(X_c(t)) \tag{4.20}$$

The real-time approximation of $I(t)$ is thus:

$$\hat{I}(t) = \hat{W}_c^T \Psi_c(X_c(t)) \quad (4.21)$$

Similar to (4.21), real-time approximation of $I(t - \tau)$ is:

$$\hat{I}(t - \tau) = \hat{W}_c^T \Psi_c(X_c(t - \tau)) \quad (4.22)$$

From (4.21) and (4.22), the temporal difference error is defined as:

$$\begin{aligned} e_c &= \hat{I}(t) - \alpha \hat{I}(t - \tau) + p_c \\ &= \hat{W}_c^T \Delta \Psi_c(t) + p_c \\ &= \tilde{W}_c^T \Delta \Psi_c(t) + p_c + W_c^T \Delta \Psi_c(t) \end{aligned} \quad (4.23)$$

where $\tilde{W}_c = W_c - \hat{W}_c$ is the weight error for critic RBFNN, $\Delta \Psi_c(t) = [\Psi_c(X_c(t)) - \alpha \Psi_c(X_c(t - \tau))]$. Define the objective function O_c as:

$$O_c = \frac{1}{2} e_c^T e_c \quad (4.24)$$

The aim is to find $\hat{W}_c$ that will minimize (4.24). Using the gradient-decent algorithm, the weight update law is designed as follows:

$$\dot{\hat{W}}_c = \gamma_c \Delta \Psi_c(t) \left[\hat{W}_c^T \Delta \Psi_c(t) + p_c \right]^T - \theta_c \gamma_c \hat{W}_c \quad (4.25)$$

where $\gamma_c = \text{diag}\{\gamma_{c_{N_l+1}} \gamma_{c_{N_l+2}} \cdots \gamma_{c_{N_l+N_f}}\}$ is a positive definite diagonal matrix, $\theta_c > 0$

is a small constant.

4.3.2 Actor design

Due to the fact that the continuous-time nonlinear model (4.6) is uncertain, accurate information about $f(q, v)$ is unavailable. Furthermore, the false data $\delta(t, q, v)$ injected into the actuators of the followers is unknown. Then, $f(q, v)$ is approximate with RBFNN as follows:

$$f(q, v) = W_{a1}^T \Psi_{a1}(X_{a1}) + \epsilon_{a1}(X_{a1}) \quad (4.26)$$

The time-varying state-dependent attack $\delta(t, q, v)$ is approximated by the action RBFNN as follows:

$$\delta(t, q, v) = W_{a2}^T \Psi_{a2}(X_{a2}) + \epsilon_{a2}(X_{a2}) \quad (4.27)$$

The real-time approximation of $f(q, v)$ and $\delta(q, v)$ are given as:

$$\hat{f}(q, v) = \hat{W}_{a1}^T \Psi_{a1}(X_{a1}) \quad (4.28)$$

$$\hat{\delta}(t, q, v) = \hat{W}_{a2}^T \Psi_{a2}(X_{a2}) \quad (4.29)$$

In the actor network, it is desired that q_f and $\dot{q}_f$ track q_f^* and $\dot{q}_f^*$ with reasonable accuracy, and minimize $\hat{I}(t)$ to the required cost value $I_d = 0$. The actor error is

defined as:

$$e_{a1} = \xi_f + \hat{I}(t) - I_d = \xi_f + \hat{W}_c^T \Psi_c \quad (4.30)$$

The actor objective function is set as:

$$O_{a1} = \frac{1}{2} e_{a1}^T e_{a1} \quad (4.31)$$

According to the gradient-decent algorithm, (4.31) can be minimized by

$$\dot{\widehat{W}}_{a1} = \gamma_{a1} \Psi_{a1}(X_{a1}) \left[\xi + \widehat{W}_c^T \Psi_c(X_c) \right]^T - \theta_{a1} \gamma_{a1} \widehat{W}_{a1} \quad (4.32)$$

where $\gamma_{a1} = \text{diag}\{\gamma_{a1N_l+1} \ \gamma_{a1N_l+2} \cdots \gamma_{a1N_l+N_f}\}$ is a positive definite diagonal matrix, $\theta_{a1} > 0$ is a small constant. Similarly, the weight update law for $\widehat{W}_{a2}$ is

$$\dot{\widehat{W}}_{a2} = \gamma_{a2} \Psi_{a2}(X_{a2}) \left[\xi + \widehat{W}_c^T \Psi_c(X_c) \right]^T - \theta_{a2} \gamma_{a2} \widehat{W}_{a2} \quad (4.33)$$

where $\gamma_{a2} = \text{diag}\{\gamma_{a2N_l+1} \ \gamma_{a2N_l+2} \cdots \gamma_{a2N_l+N_f}\}$ is a positive definite diagonal matrix, $\theta_{a2} > 0$ is a small constant.

The RL-based control laws for the nonlinear MAS (4.6) subjected to actuator

attacks are proposed as follows:

$$\begin{aligned}
\bar{u} &= K\xi_f + \underline{g}^{-1}\hat{W}_{a1}^T\Psi_{a1}(X_{a1}) + \underline{g}^{-1}\hat{W}_{a2}^T\Psi_{a2}(X_{a2})] \\
&\quad - \underline{g}^{-1}(\bar{\Omega}_{ff}^{-1}\Omega_{fl} \otimes I_d)\ddot{v}_l \\
u &= N(\beta)\bar{u} \\
\dot{\beta} &= \Lambda\xi_f^T[K\xi_f + \hat{W}_{a1}^T\Psi_{a1}(X_{a1}) + \hat{W}_{a2}^T\Psi_{a2}(X_{a2})]
\end{aligned} \tag{4.34}$$

Remark 4.3 *The essence of the Nussbaum gain $N(\beta)$ in (4.34) is to determine the control direction that is lost as a result of the actuator attacks.*

Theorem 4.1 *Consider the second-order nonlinear MAS (4.1) with actuator attacks (4.5). If assumptions (1-4) are satisfied, given the critic neural network (4.20) and actor neural networks (4.26) and (4.27), then the control law (4.34) with neural network update laws (4.32), (4.27) and (4.25) guarantees that all the closed-loop error signals are semiglobally uniformly ultimately bounded and the leader-follower AFM can be achieved. Moreover, the closed-loop error signals are trapped within a compact set defined as follows:*

$$\mathbb{S}_1 = \left\{ e_f \mid \|e_f\| \leq \sqrt{\frac{2\left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0)\right)}{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}}} \right\} \tag{4.35}$$

$$\mathbb{S}_2 = \left\{ \xi_f \mid \|\xi_f\| \leq \sqrt{2\left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0)\right)} \right\} \tag{4.36}$$

$$\mathbb{S}_3 = \left\{ \widetilde{W}_{a1} \mid \|\widetilde{W}_{a1}\| \leq \sqrt{\frac{2 \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)}{\lambda_{\min}(\gamma_{a1}^{-1})}} \right\} \quad (4.37)$$

$$\mathbb{S}_4 = \left\{ \widetilde{W}_{a2} \mid \|\widetilde{W}_{a2}\| \leq \sqrt{\frac{2 \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)}{\lambda_{\min}(\gamma_{a2}^{-1})}} \right\} \quad (4.38)$$

$$\mathbb{S}_5 = \left\{ \widetilde{W}_c \mid \|\widetilde{W}_c\| \leq \sqrt{\frac{2 \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)}{\lambda_{\min}(\gamma_c^{-1})}} \right\} \quad (4.39)$$

Proof. A candidate Lyapunov function is constructed as

$$\begin{aligned} L &= \frac{1}{2} e_f^T \bar{\Omega}_{ff}^{-1} e_f + \frac{1}{2} \xi_f^T \bar{\Omega}_{ff}^{-1} \xi_f + \frac{\Lambda}{2} \left(\widetilde{W}_{a1}^T \gamma_{a1}^{-1} \widetilde{W}_{a1} \right) \\ &\quad + \frac{\Lambda}{2} \left(\widetilde{W}_{a2}^T \gamma_{a2}^{-1} \widetilde{W}_{a2} \right) + \frac{\Lambda_c}{2} \left(\widetilde{W}_c^T \gamma_c^{-1} \widetilde{W}_c \right) \\ &\triangleq L_1 + L_2 + L_3 \end{aligned} \quad (4.40)$$

where

$$\begin{aligned} L_1 &= \frac{1}{2} e_f^T \bar{\Omega}_{ff}^{-1} e_f + \frac{1}{2} \xi_f^T \bar{\Omega}_{ff}^{-1} \xi_f \\ L_2 &= \frac{\Lambda}{2} \left(\widetilde{W}_{a1}^T \gamma_{a1}^{-1} \widetilde{W}_{a1} \right) + \frac{\Lambda}{2} \left(\widetilde{W}_{a2}^T \gamma_{a2}^{-1} \widetilde{W}_{a2} \right) \\ L_3 &= \frac{\Lambda_c}{2} \left(\widetilde{W}_c^T \gamma_c^{-1} \widetilde{W}_c \right) \end{aligned}$$

It follows that:

$$\begin{aligned}
\dot{L}_1 &= e_f^T \bar{\Omega}_{ff}^{-1} \dot{e}_f + \xi_f^T \bar{\Omega}_{ff}^{-1} \dot{\xi}_f \\
&= e_f^T \bar{\Omega}_{ff}^{-1} \dot{e}_f + \xi_f^T \bar{\Omega}_{ff}^{-1} \dot{e}_f \\
&\quad + \Lambda \xi_f^T \left\{ (f(q, v) + g(q, v)u(t) + \delta(t, q, v) + \varphi(q, v)) \right. \\
&\quad \left. + (\bar{\Omega}_{ff}^{-1} \Omega_{fl} \otimes I_d) \dot{v}_l \right\} + \dot{\beta} - \dot{\beta} \\
&= e_f^T \bar{\Omega}_{ff}^{-1} \dot{e}_f + (e_f + \Lambda \dot{e}_f)^T \bar{\Omega}_{ff}^{-1} \dot{e}_f + \Lambda \xi_f^T g(q, v)u(t) \\
&\quad + \Lambda \xi_f^T \left\{ (f(q, v) + \delta(t, q, v) + \varphi(q, v)) + \ddot{q}^* \right\} + \dot{\beta} \\
&\quad - \Lambda \xi_f^T \left[K \xi_f + \hat{W}_{a1}^T \Psi_{a1}(X_{a1}) + \hat{W}_{a2}^T \Psi_{a2}(X_{a2}) \right] \\
&= e_f^T \bar{\Omega}_{ff}^{-1} \dot{e}_f + (e_f + \Lambda \dot{e}_f)^T \bar{\Omega}_{ff}^{-1} \left(\dot{e}_f + \frac{e_f}{\Lambda} - \frac{e_f}{\Lambda} \right) \\
&\quad + \Lambda g(q, v) N(\beta) \dot{\beta} + \dot{\beta} - \Lambda \xi_f^T \widetilde{W}_{a1}^T \Psi_{a1}(X_{a1}) \\
&\quad - \Lambda \xi_f^T \widetilde{W}_{a2}^T \Psi_{a2}(X_{a2}) - \Lambda \xi_f^T K \xi_f + \Lambda \xi_f^T \ddot{q}^* \\
&\quad + \Lambda \xi_f^T \varphi(q, v) + \Lambda \xi_f^T (\epsilon_{a1} + \epsilon_{a2}) \\
&= \Lambda \xi_f^T \ddot{q}^* - \Lambda \xi_f^T K \xi_f - \frac{1}{\Lambda} e_f^T \bar{\Omega}_{ff}^{-1} e_f \\
&\quad + \frac{1}{\Lambda^2} (e_f + \Lambda \dot{e}_f)^T \bar{\Omega}_{ff}^{-1} (e_f + \Lambda \dot{e}_f) + [N(\beta) + 1] \dot{\beta} \\
&\quad - \Lambda \xi_f^T \widetilde{W}_{a1}^T \Delta \Psi_{a1}(X_{a1}) - \Lambda \xi_f^T \widetilde{W}_{a2}^T \Delta \Psi_{a2}(X_{a2}) \\
&\quad + \Lambda \xi_f^T \varphi(q, v) + \Lambda \xi_f^T \epsilon_{a1} + \Lambda \xi_f^T \epsilon_{a2}
\end{aligned} \tag{4.41}$$

$$\begin{aligned}
\dot{L}_2 &= \Lambda \left(\widetilde{W}_{a1}^T \gamma_{a1}^{-1} \dot{\widetilde{W}}_{a1} \right) + \Lambda \left(\widetilde{W}_{a2}^T \gamma_{a2}^{-1} \dot{\widetilde{W}}_{a2} \right) \\
&= \Lambda \left(\widetilde{W}_{a1}^T \gamma_{a1}^{-1} \widehat{\dot{W}}_{a1} \right) + \Lambda \left(\widetilde{W}_{a2}^T \gamma_{a2}^{-1} \widehat{\dot{W}}_{a2} \right) \\
&= -\theta_{a1} \Lambda \widetilde{W}_{a1}^T \widehat{W}_{a1} - \theta_{a2} \Lambda \widetilde{W}_{a2}^T \widehat{W}_{a2} \\
&\quad + \Lambda \xi_f^T \widetilde{W}_{a1}^T \Delta \Psi_{a1}(X_{a1}) + \Lambda \xi_f^T \widetilde{W}_{a2}^T \Delta \Psi_{a2}(X_{a2}) \\
&\quad + \Lambda \widetilde{W}_{a1}^T \Delta \Psi_{a1}(X_{a1}) [\widehat{W}_c^T \Delta \Psi_c(X_c)]^T \\
&\quad + \Lambda \widetilde{W}_{a2}^T \Delta \Psi_{a2}(X_{a2}) [\widehat{W}_c^T \Delta \Psi_c(X_c)]^T
\end{aligned} \tag{4.42}$$

$$\begin{aligned}
\dot{L}_1 + \dot{L}_2 &= \Lambda \xi_f^T \ddot{q}^* - \Lambda \xi_f^T K \xi_f - \frac{1}{\Lambda} e_f^T \bar{\Omega}_{ff}^{-1} e_f \\
&\quad + \frac{1}{\Lambda^2} \xi_f^T \bar{\Omega}_{ff}^{-1} \xi_f + [N(\beta) + 1] \dot{\beta} + \Lambda \xi_f^T \varphi(q, v) \\
&\quad + \Lambda \xi_f^T \epsilon_{a1} + \Lambda \xi_f^T \epsilon_{a2} \\
&\quad - \theta_{a1} \Lambda \widetilde{W}_{a1}^T [\widetilde{W}_{a1} + W_{a1}] - \theta_{a2} \Lambda \widetilde{W}_{a2}^T [\widetilde{W}_{a2} + W_{a2}] \\
&\quad + \Lambda \widetilde{W}_{a1}^T \Delta \Psi_{a1}(X_{a1}) [\widetilde{W}_c + W_c]^T \Delta \Psi_c(X_c)^T \\
&\quad + \Lambda \widetilde{W}_{a2}^T \Delta \Psi_{a2}(X_{a2}) [\widetilde{W}_c + W_c]^T \Delta \Psi_c(X_c)^T \\
&\leq - \left(\Lambda \lambda_{\min}(K) - 2\Lambda - \frac{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}}{\Lambda^2} \right. \\
&\quad \left. - \frac{\mu_1 + \mu_2}{2} \Lambda \right) \|\xi_f\|^2 - \frac{1}{\Lambda} \lambda_{\min}(\bar{\Omega}_{ff})^{-1} \|e_f\|^2 \\
&\quad - \Lambda \left(\frac{\theta_{a1}}{2} - \frac{\mu_1}{2} - \mu_1 \mu_2 \right) \|\widetilde{W}_{a1}\|^2 \\
&\quad - \Lambda \left(\frac{\theta_{a2}}{2} - \frac{\mu_2}{2} - \mu_1 \mu_2 \right) \|\widetilde{W}_{a2}\|^2 \\
&\quad + \frac{\mu_1 \mu_c + \mu_2 \mu_c}{2} \Lambda \|\widetilde{W}_c\|^2 + [N(\beta) + 1] \dot{\beta} + \mathcal{D}_a
\end{aligned} \tag{4.43}$$

where $\|\cdot\|$ is the Frobenius norm, $\|\Delta \Psi_{a1}\| \leq \mu_1$, $\|\Delta \Psi_{a2}\| \leq \mu_2$, $\|\Delta \Psi_c\| \leq \mu_c$, $\|W_{a1}\| \leq w_{a1}$, $\|W_{a2}\| \leq w_{a2}$, $\|W_c\| \leq w_c$, $\epsilon_{a1} \leq \bar{\epsilon}_{a1}$, $\epsilon_{a2} \leq \bar{\epsilon}_{a2}$, $\mathcal{D}_a = \frac{\Lambda \bar{\epsilon}_{a1}^2}{2} + \frac{\Lambda \bar{\epsilon}_{a2}^2}{2} + \frac{\Lambda \|\rho\|^2}{2} + \frac{\Lambda \dot{q}^{*2}}{2} +$

$\frac{\mu_1\mu_c+\mu_2\mu_c}{2}\Lambda\|w_c\|^2 + \frac{\Lambda\theta_{a1}}{2}\|w_{a1}\|^2 + \frac{\Lambda\theta_{a2}}{2}\|w_{a2}\|^2$, w_{a1} , w_{a2} , w_c , μ_1 , μ_2 , μ_c , $\bar{\epsilon}_{a1}$, and $\bar{\epsilon}_{a2}$ are

bound constants. Then, the derivative of L_3 with respect to time gives:

$$\begin{aligned}
\dot{L}_3 &= \Lambda_c \left(\widetilde{W}_c^T \gamma_c^{-1} \dot{\widetilde{W}}_c \right) = \Lambda_c \left(\widetilde{W}_c^T \gamma_c^{-1} \widehat{\dot{W}}_c \right) \\
&= -\Lambda_c \widetilde{W}_c^T \Delta \Psi_c(t) \left[\widetilde{W}_c^T \Delta \Psi_c(t) + p_c \right]^T \\
&\quad - \theta_c \Lambda_c \widetilde{W}_c^T \widetilde{W}_c \\
&= -\Lambda_c \widetilde{W}_c^T \Delta \Psi_c(t) \left[\widetilde{W}_c^T \Delta \Psi_c(t) + W_c^T \Delta \Psi_c(t) + p_c \right]^T \\
&\quad - \theta_c \Lambda_c \widetilde{W}_c^T \widetilde{W}_c - \theta_c \Lambda_c \widetilde{W}_c^T W_c \\
&\leq -\Lambda_c \left\{ \theta_c + \|\Delta \Psi_c(t)\|^2 \right\} \|\widetilde{W}_c\|^2 - \left\{ \|W_c\| \|\Delta \Psi_c(t)\|^2 \right. \\
&\quad \left. + \|\Delta \Psi_c(t)\| \|p_c\| + \theta_c \|W_c\| \right\} \Lambda_c \|\widetilde{W}_c\| \\
&\leq -\theta_c \Lambda_c \|\widetilde{W}_c\|^2 + \mathcal{D}_c \|\widetilde{W}_c\| \\
&\leq -\left\{ \theta_c \Lambda_c + \Lambda_c \|\Delta \Psi_c(t)\|^2 - \frac{1}{2} \right\} \|\widetilde{W}_c\|^2 + \frac{\mathcal{D}_c^2}{2}
\end{aligned} \tag{4.44}$$

where $\mathcal{D}_c = -\Lambda_c(w_c\mu_c^2 + \mu_c\mu_{p_c} + \theta_c w_c)$, with $\|p_c\| \leq \mu_{p_c}$. The derivative of L with respect to time is thus:

$$\begin{aligned}
\dot{L} &\leq -\left(\Lambda\lambda_{\min}(K) - 2\Lambda - \frac{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}}{\Lambda^2} \right. \\
&\quad \left. - \frac{\mu_1 + \mu_2}{2}\Lambda\right)\|\xi_f\|^2 - \frac{1}{\Lambda}\lambda_{\min}(\bar{\Omega}_{ff})^{-1}\|e_f\|^2 \\
&\quad - \Lambda\left(\frac{\theta_{a1}}{2} - \frac{\mu_1}{2} - \mu_1\mu_2\right)\|\widetilde{W}_{a1}\|^2 \\
&\quad - \Lambda\left(\frac{\theta_{a2}}{2} - \frac{\mu_2}{2} - \mu_1\mu_2\right)\|\widetilde{W}_{a2}\|^2 \\
&\quad - \left(\Lambda_c\theta_c + \Lambda_c\mu_c^2 - \frac{1}{2} - \frac{\mu_1\mu_c + \mu_2\mu_c}{2}\Lambda\right)\|\widetilde{W}_c\|^2 \\
&\quad + [N(\beta) + 1]\dot{\beta} + \mathcal{D}_a + \frac{\mathcal{D}_c^2}{2} \\
&\leq -\mathcal{D}_1L + [N(\beta) + 1]\dot{\beta} + \mathcal{D}_2
\end{aligned} \tag{4.45}$$

where $\mathcal{D}_2 = \mathcal{D}_a + \frac{\mathcal{D}_c^2}{2}$,

$$\begin{aligned}
\mathcal{D}_1 &= \min\left\{\left(\Lambda\lambda_{\min}(K) - 2\Lambda - \frac{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}}{\Lambda^2} \right. \right. \\
&\quad \left. \left. - \frac{\mu_1 + \mu_2}{2}\Lambda\right), \frac{1}{\Lambda}\lambda_{\min}(\bar{\Omega}_{ff})^{-1}, \right. \\
&\quad \Lambda\left(\frac{\theta_{a1}}{2} - \frac{\mu_1}{2} - \mu_1\mu_2\right), \Lambda\left(\frac{\theta_{a2}}{2} - \frac{\mu_2}{2} - \mu_1\mu_2\right), \\
&\quad \left.\left(\Lambda_c\theta_c + \Lambda_c\mu_c^2 - \frac{1}{2} - \frac{\mu_1\mu_c + \mu_2\mu_c}{2}\Lambda\right)\right\}
\end{aligned} \tag{4.46}$$

To guarantee that $\mathcal{D}_1 > 0$, the parameters Λ_c , Λ , θ_{a1} , θ_{a2} , θ_c , and the matrix K

are selected such that the following conditions are satisfied:

$$\begin{aligned}
& \Lambda_c > 0; \quad \Lambda > 0; \quad \theta_{a1} > 0; \quad \theta_{a2} > 0; \quad \theta_c > 0 \\
& \lambda_{\min}(K) > 0; \quad \lambda_{\min}(\bar{\Omega}_{ff})^{-1} > 0; \quad \lambda_{\min}(K) > \Lambda \\
& \left(\Lambda \lambda_{\min}(K) - 2\Lambda - \frac{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}}{\Lambda^2} - \frac{\mu_1 + \mu_2}{2} \Lambda \right) > 0 \\
& \left(\frac{\theta_{a1}}{2} - \frac{\mu_1}{2} - \mu_1 \mu_2 \right) > 0 \\
& \left(\frac{\theta_{a2}}{2} - \frac{\mu_2}{2} - \mu_1 \mu_2 \right) > 0 \\
& \left(\Lambda_c \theta_c + \Lambda_c \mu_c^2 - \frac{1}{2} - \frac{\mu_1 \mu_c + \mu_2 \mu_c}{2} \Lambda \right) > 0
\end{aligned}$$

Having $\mathcal{D}_1 > 0$ will guarantee the convergence of the closed-loop tracking errors into the compact sets (4.35)-(4.39).

By multiplying both sides of (4.46) by $e^{-\mathcal{D}_1 t}$, it can be rewritten as follows:

$$e^{\mathcal{D}_1 t} \dot{L} \leq -\mathcal{D}_1 L e^{\mathcal{D}_1 t} + [N(\beta) + 1] \dot{\beta} e^{\mathcal{D}_1 t} + \mathcal{D}_2 e^{\mathcal{D}_1 t} \quad (4.47)$$

The inequality (4.47) further yields:

$$\frac{d}{dt}(e^{\mathcal{D}_1 t} L(t)) \leq [N(\beta) + 1] \dot{\beta} e^{\mathcal{D}_1 t} + \mathcal{D}_2 e^{\mathcal{D}_1 t} \quad (4.48)$$

Computing the integral of (4.48) over $[0, t)$, we get:

$$\begin{aligned}
L(t) &\leq \frac{\mathcal{D}_2}{\mathcal{D}_1} + L(0)e^{-\mathcal{D}_1 t} \\
&\quad + e^{-\mathcal{D}_1 t} \int_0^t [N(\beta) + 1] \dot{\beta} e^{\mathcal{D}_1 \iota} d\iota
\end{aligned} \tag{4.49}$$

Based on Lemma 4.1, $\beta(t)$, $\int_0^t [N(\beta) + 1] \dot{\beta} e^{\mathcal{D}_1 \iota} d\iota$ and $L(t)$ are bounded. It follows that all the closed-loop error signals, ξ_f , $\widetilde{W}_{a1}$, $\widetilde{W}_{a2}$ and $\widetilde{W}_c$ are bounded. Let $\mathcal{D}_3$ be the upper-bound of $e^{-\mathcal{D}_1 t} \int_0^t [N(\beta) + 1] \dot{\beta} e^{\mathcal{D}_1 \iota} d\iota$, then (4.49) can be simplified as:

$$\begin{aligned}
L(t) &\leq \frac{\mathcal{D}_2}{\mathcal{D}_1} + L(0)e^{-\mathcal{D}_1 t} + \mathcal{D}_3 \\
&\leq \frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0)
\end{aligned} \tag{4.50}$$

From the Lyapunov function (4.40) and the inequality (4.50), it can be obtain that:

$$\|e_f\| \leq \sqrt{\frac{2}{\lambda_{\min}(\bar{\Omega}_{ff})^{-1}} \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)} \tag{4.51}$$

$$\|\xi_f\| \leq \sqrt{2 \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)} \tag{4.52}$$

$$\|\widetilde{W}_{a1}\| \leq \sqrt{\frac{2}{\lambda_{\min}(\gamma_{a1}^{-1})} \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)} \tag{4.53}$$

$$\|\widetilde{W}_{a2}\| \leq \sqrt{\frac{2}{\lambda_{\min}(\gamma_{a2}^{-1})} \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)} \tag{4.54}$$

$$\|\widetilde{W}_c\| \leq \sqrt{\frac{2}{\lambda_{\min}(\gamma_c^{-1})} \left(\frac{\mathcal{D}_2}{\mathcal{D}_1} + \mathcal{D}_3 + L(0) \right)} \quad (4.55)$$

From (4.46), it can be seen that $\mathcal{D}_1$ can be manipulated by varying the parameters Λ_c , Λ , θ_{a1} , θ_{a2} , θ_c , and the matrix K . Therefore, the tracking errors ξ_f , $\widetilde{W}_{a1}$, $\widetilde{W}_{a2}$ and $\widetilde{W}_c$ can be made arbitrarily small by carefully adjusting the parameters Λ_c , Λ , θ_{a1} , θ_{a2} , θ_c , and the matrix K to increase $\mathcal{D}_1$. Large enough $\mathcal{D}_1$ makes the ratio $\mathcal{D}_2/\mathcal{D}_1$ smaller and subsequently reduces the error signals (4.51)-(4.55). ■

Remark 4.4 *The problem of AFM control of MAS has been studied in [21]-[68]. However, the authors did not consider the possibility of cyber-attacks on the MAS. These attacks can disrupt the exchange of information among the agents and subsequently leads to the collapse of the MAS. In this article, for the first time, the secured AFM control of MAS is investigated. The distributed control laws (4.34) for the MAS (4.6) have attack estimation and suppression mechanisms via the actor RBFNN contrary to [21]-[68].*

Remark 4.5 *In contrast to [21]-[68], the distributed control laws (4.34) are formulated to minimize the long-term performance index (4.16) using the actor-critic RL architecture.*

Remark 4.6 *The authors in [68] proposed an AFM control of uncertain second-order nonlinear MAS similar to this article. Nonetheless, the MAS in [68] belongs to the class of nonlinear systems that can be expressed in linear-in-the-parameter form. The uncertain parameters were estimated using adaptive laws. In contrast, the nonlinear*

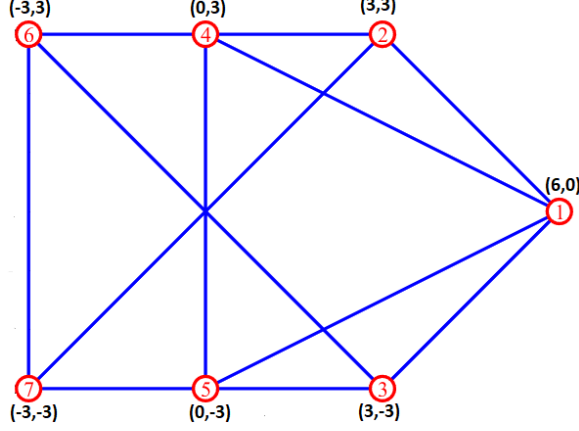


Figure 4.1: The nominal formation showing the inter-agents communications.

MAS in this paper is not required to be in the linear-in-the-parameter form since actor RBFNN are employed to estimate them as shown in (4.34).

4.4 Simulation results

This section presents a numerical simulation to demonstrate the validity of the actor-critic RL scheme for achieving the AFM control of the leader-follower MAS subjected to actuator attacks. The MAS under this study consists of three leaders and four followers. The nominal formation of the agents is depicted in Fig. 4.1. The nominal configuration of the agents is given in (4.56), where $r_1 = [6 \ 0]$, $r_2 = [3 \ 3]$, $r_3 = [3 \ -3]$, $r_4 = [0 \ 3]$, $r_5 = [0 \ -3]$, $r_6 = [-3 \ 3]$ and $r_7 = [-3 \ -3]$. Using the method presented in [49], the matrix of the stresses among the agents is computed as (4.57).

$$r = \begin{bmatrix} r_1^T & r_2^T & r_3^T & r_4^T & r_5^T & r_6^T & r_7^T \end{bmatrix}^T \quad (4.56)$$

$$\Omega = \begin{bmatrix} -0.2741 & 0.2741 & 0.2741 & -0.1370 & -0.1370 & 0 & 0 \\ 0.2741 & -0.6852 & 0 & 0.5482 & 0 & 0 & -0.1370 \\ 0.2741 & 0 & -0.6852 & 0 & 0.5482 & -0.1370 & 0 \\ -0.1370 & 0.5482 & 0 & -0.7537 & 0.0685 & 0.2741 & 0 \\ -0.1370 & 0 & 0.5482 & 0.0685 & -0.7537 & 0 & 0.2741 \\ 0 & 0 & -0.1370 & 0.2741 & 0 & -0.2741 & 0.1370 \\ 0 & -0.1370 & 0 & 0 & 0.2741 & 0.1370 & -0.2741 \end{bmatrix} \quad (4.57)$$

The nonlinear dynamic equations of the four followers are given as:

$$\begin{aligned} \dot{q}_i &= v_i \\ \dot{v}_i &= \begin{bmatrix} q_{i1}^2 + 2v_{i1}\cos(q_{i1})v_{i1} \\ q_{i2}^2 + 2v_{i2}\cos(q_{i2})v_{i2} \end{bmatrix} + \begin{bmatrix} u_{i1}^c \\ u_{i2}^c \end{bmatrix} + \begin{bmatrix} \varphi_{i1}(q_{i1}, v_{i1}) \\ \varphi_{i2}(q_{i2}, v_{i2}) \end{bmatrix} \end{aligned} \quad (4.58)$$

where $\varphi_{i1}(q_{i1}, v_{i1}) = 0.22v_{i1}\cos(0.7q_{i1}t)$, $\varphi_{i2}(q_{i2}, v_{i2}) = 0.43v_{i2}\cos(2q_{i2}t)$.

The MAS encountered two false data injection attacks on the actuators of the followers. The first attack targeted the actuator of the fourth follower for $40s \leq t < 80s$, whereas the second attack targeted the actuator of the first follower for $210s \leq t < 240s$. The comprised actuators of the fourth and the first followers as a result of the attacks are expressed as $u_{41}^c = (1.2 + 0.45\cos(q_{41})u_{41} - 62\cos(20\pi t)$, $u_{42}^c = (1.2 + 0.45\cos(q_{42})u_{42} - 62\sin(20\pi t)$, $u_{11}^c = (1.2 + 0.45\sin(q_{11})u_{11} + 62\cos(20\pi t)$, and $u_{12}^c = (1.2 + 0.45\sin(q_{12})u_{12} + 62\cos(20\pi t)$,

The initial positions of the agents are set as $r_1(0) = [6 \ 0]^T$, $r_2(0) = [3 \ 3]^T$, $r_3(0) = [3 \ -3]^T$, $r_4(0) = [1 \ 4]^T$, $r_5(0) = [0 \ -2]^T$, $r_6(0) = [-4 \ 4]^T$, $r_7(0) = [-5 \ -4]^T$. The integral reinforcement interval is set as $\tau = 0.2$, The gains of the controllers are carefully chosen as, $\Lambda = 2$, $K = 25I$. The parameters of the critic neural network are set as $\gamma_c = 0.2I$, $\theta_c = 0.1$, the width of the Gaussian function $\eta_c = 3$, the center of

the receptive field c_c is chosen between $-2, -1.5, -1, \dots, 2$, number of hidden nodes is chosen as 10, the initial critic weight vector is selected as $W_{c_0} = [0, 0, \dots, 0]^T$. The parameters of the actor neural network are set as $\gamma_{a1} = 6I$, $\gamma_{a2} = 6I$, $\theta_{a1} = 0.1$, $\theta_{a2} = 0.1$, the width of the Gaussian function $\eta_{a1} = 3$, $\eta_{a2} = 3$, the center of the receptive field c_{a1} and c_{a2} are chosen between $-1, -0.5, \dots, 2$, number of hidden nodes for the first and second actor neural networks are chosen as 30 and 15, respectively, the initial actor weight vectors are selected as $W_{a1_0} = [0.2, 0.2, \dots, 0.2]^T$ and $W_{a2_0} = [0.2, 0.2, \dots, 0.2]^T$.

It is worth noting that the AFM control method presented in [68] did not take into account the impact of cyber-attacks on the MAS. In order to demonstrate the failure of the AFM control approach proposed in [68] in the event of false-data injection attacks, it is applied to the MAS (4.1). Fig. 4.2 depicts the distorted trajectories of the followers q_4 and q_7 for $40s \leq t < 80s$ and $210s \leq t < 240s$, respectively as a result of the attacks on their actuators. In practical applications, the attacks will lead to the collision of the followers. Therefore, the AFM control approach proposed in [68] is not suitable for application in adversarial environments. As shown in Fig. 4.3, the trajectories of the agents are smooth because the proposed control technique is able to detect and mitigate the attacks on the actuators of the followers unlike [68]. The AFM tracking errors of the followers under the proposed scheme are depicted in Fig. 4.4. The responses of the RL-based distributed control protocols that mitigate the actuator attacks and maintain the AFM of the leader-follower system MAS is depicted in Fig. 4.5. The evolution of the norm-2 of the RBFNN weight vectors of the two actors and the critic networks is presented in Fig. 4.6, Fig. 4.7, and Fig.

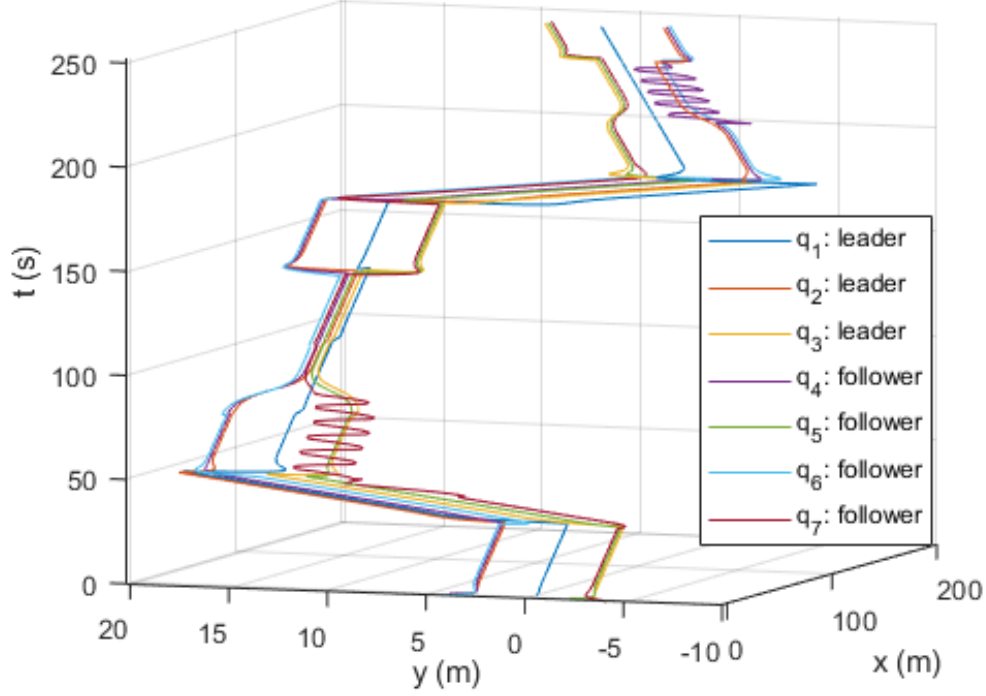


Figure 4.2: Effects of attacks on the trajectories of the leader-follower MAS under the control input proposed in (Zhi and Chen 2021).

4.8, respectively. It can be observed that the weights are automatically adjusted to adapt to the new collective maneuvers of the agents. In order, to show the accuracy of the RBFNN in approximating the unknown function $f(q, v)$, the approximation errors $(f(q, v) - \hat{f}(q, v))$ are presented in Fig. 4.9. It can be observed that the approximation errors vary between -0.04 and 0.04 during the steady states. Fig. 4.10 shows the collective maneuvering of the formation of the agents such as translation, collinearity, scaling, shearing, and rotation.

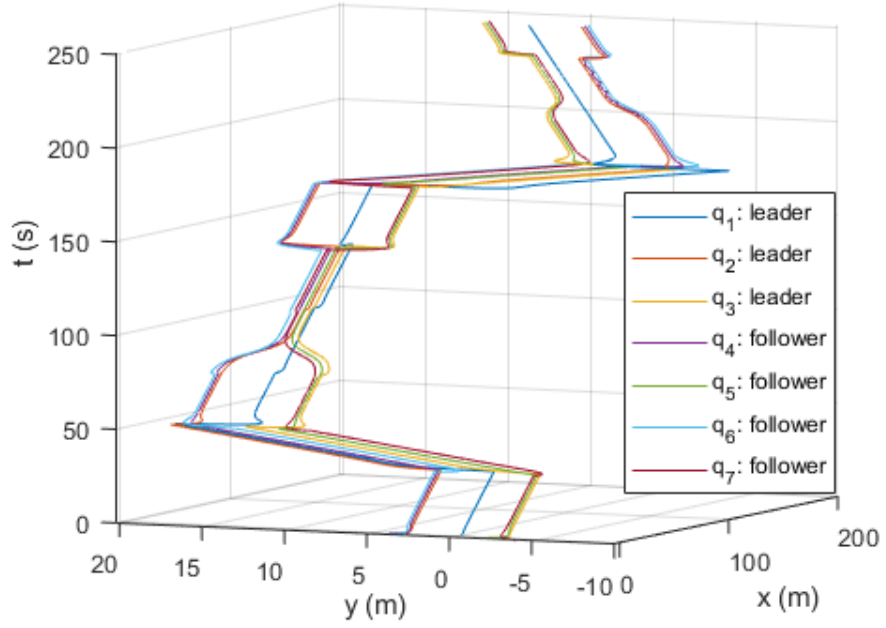


Figure 4.3: The trajectories of the leader-follower MAS under the attack mitigation control protocol (4.34).

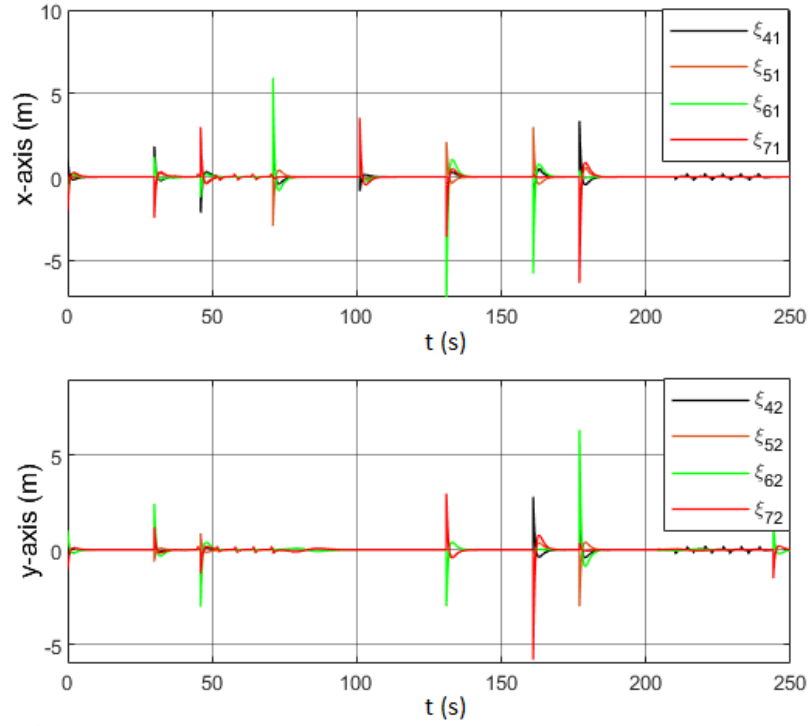


Figure 4.4: The AFM tracking errors between the followers and the leaders under the proposed scheme.

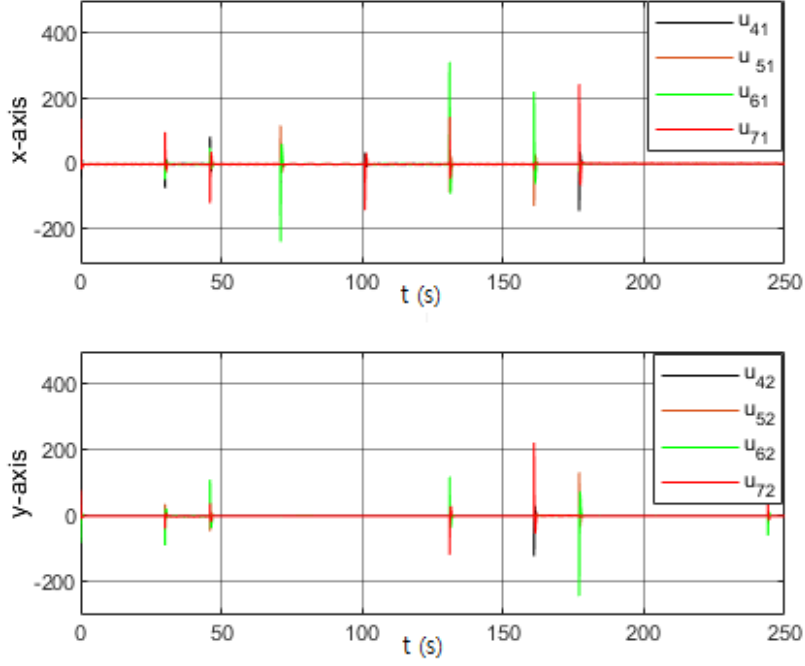


Figure 4.5: The RL control signals of the followers that mitigate actuator attacks and maintain the AFM.

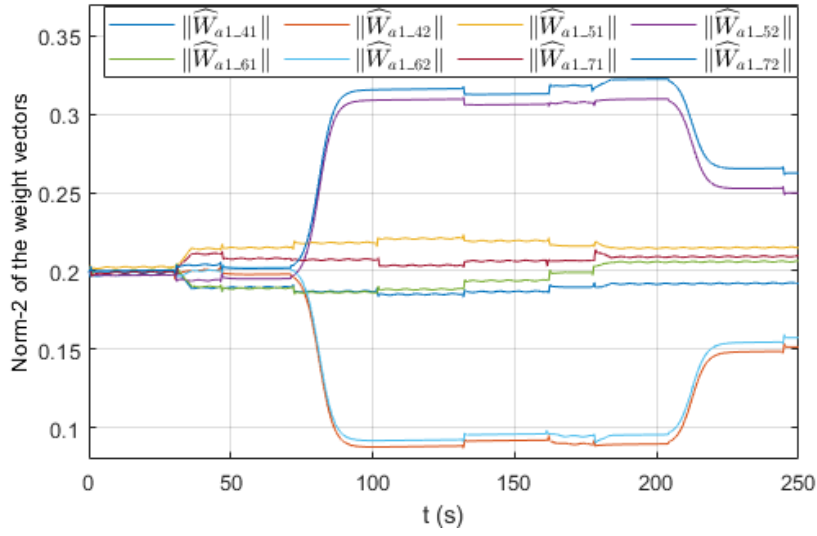


Figure 4.6: The update of the actor weights for the RBFNN approximation of $f(q, v)$.

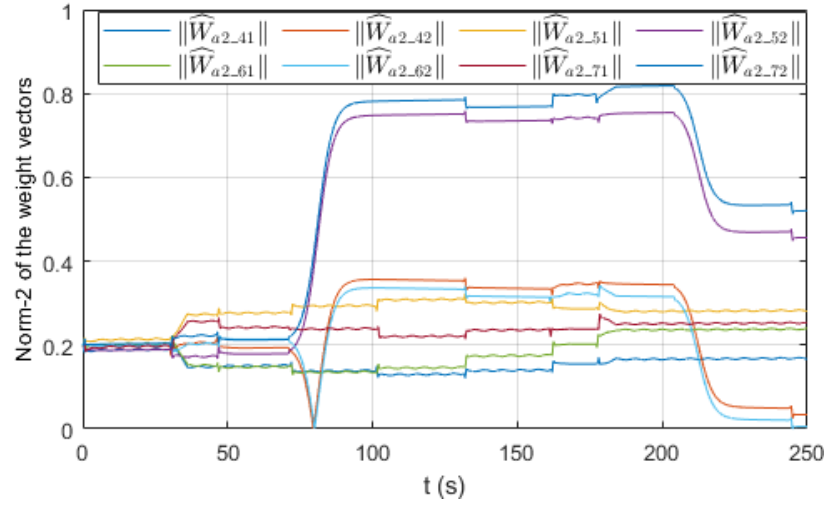


Figure 4.7: The update of the actor weights for the RBFNN approximation of $\delta(q, v)$.

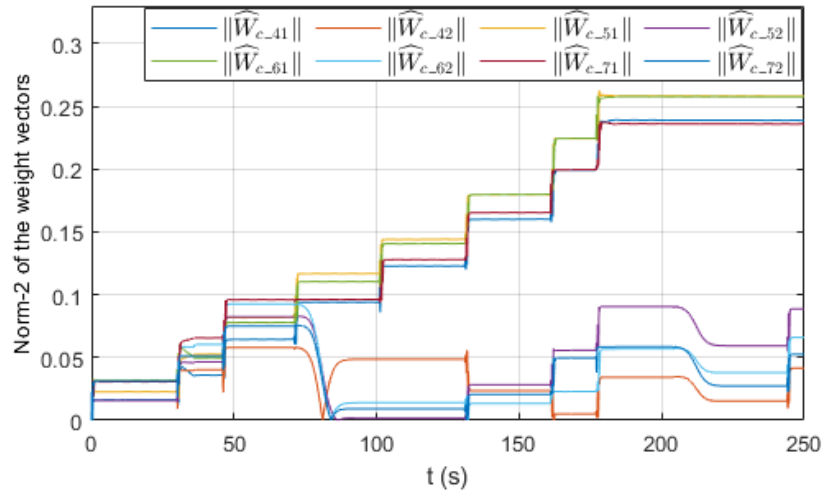


Figure 4.8: The update of the critic weights for the RBFNN approximation of $I(t)$.

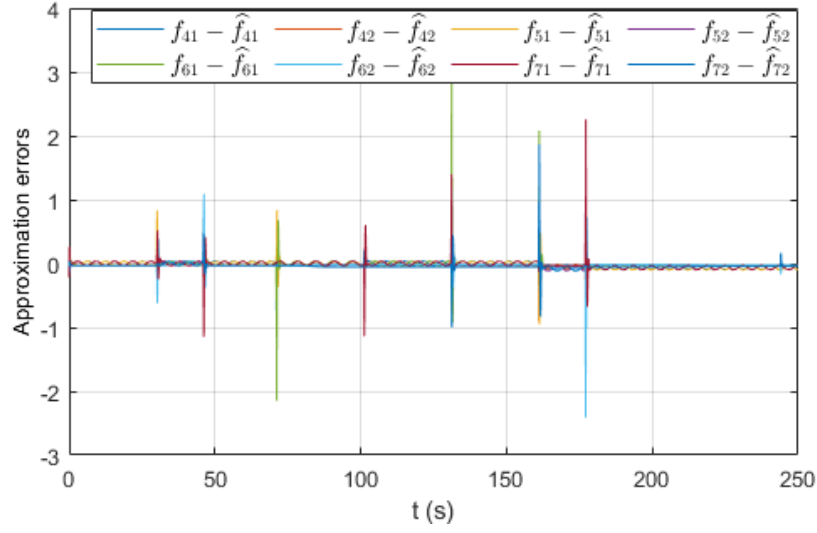


Figure 4.9: The approximation errors between $f(q.v)$ and $\hat{f}(q,v)$.

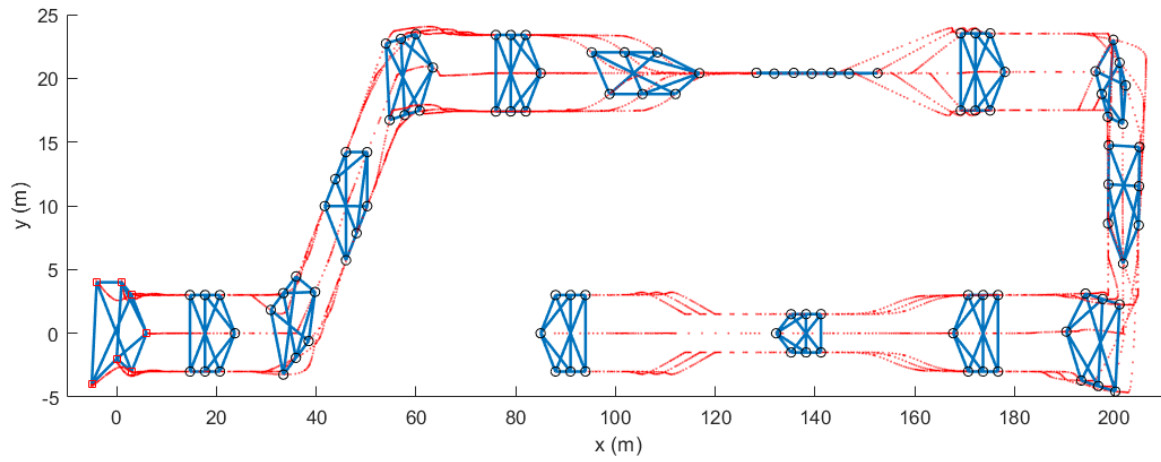


Figure 4.10: Trajectories showing various affine formation maneuvers of the leaders and followers.

CHAPTER 5

SECURE REINFORCEMENT LEARNING-BASED FINITE-TIME AFFINE FORMATION MANEUVER CONTROL OF MULTI-AGENT SYSTEMS SUBJECTED TO DECEPTION ATTACKS AND BACKLASH-LIKE HYSTERESIS

This chapter provides a leader-follower affine formation maneuver (AFM) problem for uncertain nonlinear multi-agent systems, in which some complex and challenging issues, such as unknown external disturbances, unknown input backlash-like hysteresis, and sensor deception attacks are addressed with the aid of the actor-critic learning control based on neural networks. The main control objective is to steer the group of followers to track the maneuvers of the group of leaders within a finite time. The control protocol is built using a distributed fast terminal sliding mode surface equipped with a reaching law to facilitate finite-time convergence. The actor neural network is handed the responsibility of estimating the uncertain nonlinear dynamics, unknown input backlash-like hysteresis, and deception attack signals and compensating them through the control protocol. Meanwhile, the critic neural network evaluates the operation of the control protocol. Lyapunov stability theorem is deployed to prove that the closed-loop system is bounded in the vicinity of zero in finite time. Simulation

results are provided to demonstrate the effectiveness of the proposed control scheme.

5.1 Introduction

In real-world cyber-physical systems, the exchange of information between mobile agents during formation maneuvering and the networked control systems is exposed to malicious cyber-attacks. Cyber-attacks unleashed by rivals are commonly perpetrated in three ways: deception attacks [119], denial of service attacks [120], and replay attacks [121]. The aim of the attackers is to obliterate the integrity of the transmitted information, and ultimately degrade the formation control performance. Therefore, it is crucial to investigate the resilient formation control problems in the presence of cyber-attacks. In this article, we are concerned with deception attacks. Resilient formation control against deception attacks has been proposed for a group of quadrotor unmanned aerial vehicles [122], [123], [124], and swarm of mobile robots [28], [125], [126]. The fact that these resilient formation control techniques mainly focus on time-varying formation tracking of a single leader or virtual leader with multiple followers, we are motivated to investigate the resilient AFM control of multiple leaders with multiple followers such that the resulting formation can undergo several affine maneuvers at the same time.

The presence of nonstructured nonlinearities (such as saturation, deadzone, and backlash-like hysteresis) can potentially degrade the control performance of a dynamic system. Backlash-like hysteresis is widely present in the actuators of mechanical systems and tackling it requires urgent attention. Several studies have successfully addressed this problem [127], [128]. However, in the existing literature on AFM con-

trol, the negative effects of backlash-like hysteresis have not been studied.

So far, the existing AFM control problems are based on asymptotic convergence, that is, the followers converge to the maneuvers of the followers as the time approaches infinity. Nonetheless, in practical application, especially in adversarial environments, a short finite time is required to mitigate the cyber-attacks launched on the agents. Thus, finite-time AFM control deserves special attention. The authors in [129] provided a comprehensive overview of the applications and advantages of finite-time control methods.

Considering the foregoing discussion, this chapter presents finite-time AFM control of nonlinear multi-agent systems subjected to deception attacks and backlash-like hysteresis.

1. Contrary to existing AFM control methods which are mainly proportional-integral, proportional-derivative, or proportional-integral-derivative with asymptotic stability, a distributed fast terminal sliding-mode-like control is formulated so that reaching laws can be invoked to enhance the robustness and guarantee finite-time stability. Thus, the control approach developed in this chapter can speed up the convergence rate of the leader-follower formation maneuvering.
2. Different from the usual AFM control of MASs, the considered nonlinear MAS model encountered external disturbances, sensor deception attacks, input backlash-like hysteresis, and unknown input gains. No doubt, these problems add difficulties to the control design.

3. The actor-critic reinforcement learning is used to estimate the uncertain nonlinear dynamics, unknown input backlash-like hysteresis, and deception attack signals and compensate them through the distributed control law.

5.2 Problem Formulation

5.2.1 Dynamic models of the agents

This article considers leader-follower networked systems comprising N agents, where N_l of the agents are the leaders while $N_f = N - N_l$ of the agents are the followers. The motions of the agents are modeled by the following nonlinear equations:

$$\ddot{q}_i + C_i(q_i, \dot{q}_i) + H_i(q_i) = B_i(\tau_i) + \tau_{di} \quad (5.1)$$

where $q_i \in \mathbb{R}^n$ is the position vector of the i th agent, $\tau_i \in \mathbb{R}^n$ control torque, $\tau_{di} \in \mathbb{R}^n$ is the external disturbance, $C_i(q_i, \dot{q}_i) \in \mathbb{R}^{n \times n}$ is a vector of nonlinear functions, $H_i(q_i) \in \mathbb{R}^{n \times n}$ is a diagonal matrix. e.

5.2.2 Backlash-like Hysteresis

The backlash-like hysteresis is modeled with a first-order differential equation as follows:

$$\frac{dB_i(\tau_i)}{dt} = c_{1i} \left| \frac{d\tau}{dt} \right| [c_{2i}\tau_i - B_i(\tau_i)] + c_{3i} \frac{d\tau}{dt} \quad (5.2)$$

where $c_{1i} \in \mathbb{R}^{n \times n}$, $c_{2i} \in \mathbb{R}^{n \times n}$ and $c_{3i} \in \mathbb{R}^{n \times n}$ are positive-definite constant diagonal matrices. The solution of the differential equation (5.2) yields:

$$B_i(\tau_i) = c_{2i}\tau_i + \partial_i(\tau_i) \quad (5.3)$$

$$\begin{aligned} \partial_i(\tau_i) &= (B_{0i} - c_{2i}\tau_{0i}) e^{-c_{1i}(\tau_i - \tau_{0i}) \operatorname{sgn}(\dot{\tau}_i)} \\ &\quad + e^{-c_{1i}\tau_i \operatorname{sgn}(\dot{\tau}_i)} \int_{\tau_{0i}}^{\tau_i} (c_{3i} - c_{2i}) e^{c_{1i}\beta_j \operatorname{sgn}(\dot{\tau}_i)} d\beta_j \end{aligned} \quad (5.4)$$

where B_{0i} and τ_{0i} are the initial values of B_i and τ_i , respectively.

Assumption 5.1 *Given that the constant matrices c_{1i} , c_{2i} and c_{3i} are unknown and bounded, it follows that the function $\partial_i(\tau_i)$ is unknown and bounded.*

5.2.3 Sensor Attack model

The deceptive data injected into to sensors of cyber-physical systems tend to corrupt the measured state signals. The sensor attack signals are modeled state-dependent variables as:

$$\check{q}_i = q_i + \varpi_i(t)q_i = \pi_i(t)q_i \quad (5.5)$$

where $\check{q}_i \in \mathbb{R}^n$ is the contaminated state signal, $\varpi_i(t) \in \mathbb{R}^{n \times n}$ is an unknown time-varying attack weight, $\varpi_i(t)q_i$ is the state-dependent malicious signal, $\pi_i = (1 + \varpi_i(t)) \in \mathbb{R}^{n \times n}$.

Remark 5.1 *In the cyber transmission of data for feedback control, the state signals measured by the sensors are susceptible to attacks from adversaries. When the sensors*

are injected with harmful signals, the corrupted state variables $\check{q}_i$ will be utilized for control design because the healthy state variables q_i are no longer available.

For the purpose of control design, the Euler-Langrange system (5.1) will be converted to the compromised form. The time-derivative of (5.5) gives:

$$\dot{\check{q}}_i = \dot{\pi}_i \check{q}_i + \pi_i \dot{q}_i \quad (5.6)$$

$$\ddot{\check{q}}_i = \ddot{\pi}_i \check{q}_i + 2\dot{\pi}_i \dot{q}_i + \pi_i \ddot{q}_i \quad (5.7)$$

Substituting $\check{q}_i = \frac{\check{q}_i}{\pi_i}$, $\dot{q}_i = \frac{1}{\pi_i} \left[\dot{q}_i - \dot{\pi}_i \frac{\check{q}_i}{\pi_i} \right]$, (5.1) and (5.3) into (5.7) gives:

$$\begin{aligned} \ddot{\check{q}}_i &= \ddot{\pi}_i \frac{\check{q}_i}{\pi_i} + 2 \frac{\dot{\pi}_i}{\pi_i} \left[\dot{q}_i - \dot{\pi}_i \frac{\check{q}_i}{\pi_i} \right] - \pi_i C_i (\check{q}_i, \dot{\check{q}}_i) \\ &\quad - \pi_i H_i(q_i) + \pi_i c_{2i} \tau_i + \pi_i \partial_i + \pi_i \tau_{di} \end{aligned} \quad (5.8)$$

The compromised dynamics of an i th agent is thus:

$$\ddot{\check{q}}_i = F_i + G_i \tau_i \quad (5.9)$$

where

$$\begin{aligned} F_i &= \ddot{\pi}_i \frac{\check{q}_i}{\pi_i} + 2 \frac{\dot{\pi}_i}{\pi_i} \left[\dot{q}_i - \dot{\pi}_i \frac{\check{q}_i}{\pi_i} \right] - \pi_i C_i (\check{q}_i, \dot{\check{q}}_i) \\ &\quad - \pi_i H_i(q_i) + \pi_i \partial_i + \pi_i \tau_{di} \\ G_i &= \pi_i c_{2i} \end{aligned}$$

Lemma 5.1 [130] For any variable y , an arbitrary constant h , $0 < \gamma < 1$, and

$0 < \varrho < 1$, the following relationship is satisfied:

$$-y(y+h)^\gamma \leq -\frac{1-\varrho}{1+\gamma}y^{1+\gamma} + \frac{\bar{h}}{1+\gamma} \quad (5.10)$$

where $\bar{h} = h^{1+\gamma} + \left(\frac{h}{1-(1-\varrho)^{1/(1+\gamma)}}\right)^{1+\gamma} + \left(\frac{h(1-\varrho)^{1/(1+\gamma)}}{1-(1-\varrho)^{1/(1+\gamma)}}\right)^{1+\gamma}$.

Lemma 5.2 [131] *For any real number $a_1 > 0$, $a_2 > 0$, $0 < \gamma < 1$, a Lyapunov candidate function of finite-time stability satisfies $\dot{L} + a_1 L + a_2 L^\gamma \leq 0$, with a settling time given by $T_s \leq t_0 + \left[\frac{1}{a_1(1-\gamma)}\right] \ln \left[\frac{a_1 L^{1-\gamma}(t_0) + a_2}{a_2}\right]$*

5.3 Control design

This section presents a secure reinforcement learning-based control law τ_i to address to achieve a secure leader-follower AFM for the MAS (5.9) under sensor deception attacks, backlash-like hysteresis, and unknown input gains. It is desired that:

$$\check{q}_f \longrightarrow \check{q}_f^*(t) = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_2) \check{q}_l^* \text{ as } t \longrightarrow T_s \quad (5.11)$$

where T_s is a finite settling time.

The distributed fast terminal sliding mode surfaces of the AFM tracking errors of the followers with respect to the leaders are given by:

$$\begin{aligned} \zeta_i = & \sum_{j=1}^N \omega_{ij}(\check{\dot{q}}_i - \check{\dot{q}}_j) + \Lambda \sum_{j=1}^N \omega_{ij}(\check{q}_i - \check{q}_j) \\ & + \rho \sum_{j=1}^N \omega_{ij}(\check{q}_i - q_j)^{\frac{b_2}{b_1}}, \quad i \in \mathcal{V}_f \end{aligned} \quad (5.12)$$

where $\Lambda > 0$ and $\rho > 0$ are design constants, $b_1 > 0$ and $b_2 > 0$ are selected such that $0 < b_1 < 1$ and $b_2 = \frac{2b_1+1}{b_1+1}$

Let $z_i = \sum_{j=1}^N \omega_{ij}(\check{q}_i - \check{q}_j)$, $\dot{z}_i = \sum_{j=1}^N \omega_{ij}(\dot{\check{q}}_i - \dot{\check{q}}_j)$, and $z_i^{\frac{b_2}{b_1}} = \sum_{j=1}^N \omega_{ij}(\check{q}_i - \check{q}_j)^{\frac{b_2}{b_1}}$

Equation (5.12) can be rewritten as:

$$\zeta_i = \dot{z}_i + \Lambda z_i + \rho z_i^{\frac{b_2}{b_1}}, \quad i \in \mathcal{V}_f \quad (5.13)$$

The compact form of z_i and $\dot{z}_i$ are given as [21]:

$$z_f = \bar{\Omega}_{ff} \check{q}_f + \bar{\Omega}_{fl} \check{q}_l^* = \bar{\Omega}_{ff}(\check{q}_f - q_f^*) \quad (5.14)$$

$$z_f^{\frac{b_2}{b_1}} = (\bar{\Omega}_{ff} \check{q}_f + \bar{\Omega}_{fl} \check{q}_l^*)^{\frac{b_2}{b_1}} = (\bar{\Omega}_{ff}(\check{q}_f - q_f^*))^{\frac{b_2}{b_1}} \quad (5.15)$$

$$\dot{z}_f = \bar{\Omega}_{ff} \dot{\check{q}}_f + \bar{\Omega}_{fl} \dot{\check{q}}_l^* = \bar{\Omega}_{ff}(\dot{\check{q}}_f - \dot{q}_f^*) \quad (5.16)$$

where $z_f = [z_{N_l+1}^T \ z_{N_l+1}^T \cdots z_{N_l+N_f}^T]^T$, and $q_f^* = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_2) \check{q}_l^*$

The compact form of (5.13) is thus:

$$\dot{\zeta}_f = \dot{z}_f + \Lambda z_f + \rho z_f^{\frac{b_2}{b_1}} \quad (5.17)$$

The derivative of (5.17) with respect to time gives:

$$\dot{\zeta}_f = \bar{M}_{ff} \left\{ F + G\tau - \check{q}_f^* \right\} + \Lambda \dot{z}_f + \rho \frac{b_2}{b_1} \text{diag}\{\dot{z}_f\} z_f^{\frac{b_2}{b_1}-1} \quad (5.18)$$

where $\zeta_f = [\zeta_{N_l+1}^T \ \zeta_{N_l+1}^T \cdots \zeta_{N_l+N_f}^T]^T$, $z_f = [z_{N_l+1}^T \ z_{N_l+1}^T \cdots z_{N_l+N_f}^T]^T$,
 $F = [F_{N_l+1}^T \ F_{N_l+2}^T \cdots F_{N_l+N_f}^T]^T$, $\tau = [\tau_{N_l+1}^T \ \tau_{N_l+2}^T \cdots \tau_{N_l+N_f}^T]^T$, $G =$
 $\text{diag}\{G_{N_l+1}^T \ G_{N_l+2}^T \cdots G_{N_l+N_f}^T\}$,

5.3.1 Critic design

To assess the tracking performance of the AFM control of the agents, a long-term performance measure is given as:

$$\Upsilon(t) = \int_t^\infty \beta^{\frac{-\vartheta+t}{\delta}} b(\zeta(\vartheta)) d\vartheta \quad (5.19)$$

where $\Upsilon(t) = [\Upsilon_{N_l+1}^T \ \Upsilon_{N_l+2}^T \ \dots \ \Upsilon_{N_l+N_f}^T]^T$ is the long term performance measure, $\beta \in (0, 1)$ is the discount factor, $\delta > 0$ is the integral reinforcement interval, $b(\zeta) = [b(\zeta_{N_l+1})^T \ b(\zeta_{N_l+2})^T \ \dots \ b(\zeta_{N_l+N_f})^T]^T$ is the binary utility function expressed as:

$$b(\zeta_i(t)) = \begin{cases} 0 & \text{if } \zeta_i(t)^2 \leq h_i \\ 1 & \text{if } \zeta_i(t)^2 > h_i \end{cases} \quad \forall \zeta \in [t-T, t) \quad (5.20)$$

where $h_i > 0$ is a prescribed threshold. When the AFM tracking error ζ_i is within the prescribed threshold, $b(\zeta_i(t)) = 0$; otherwise, $b(\zeta_i(t)) = 1$. tracking performance. The function (5.20) is widely employed for reinforcement learning in the discrete-time domain [132], [133].

To derive the continuous-time Bellman error, the relationship between $\Upsilon(t)$ and $\Upsilon(t - \delta)$ can be expressed as:

$$\begin{aligned} \Upsilon(t - \delta) &= \int_{t-\delta}^\infty \beta^{\frac{-\vartheta+t-\delta}{\delta}} b(\zeta(\vartheta)) d\vartheta \\ &= \beta^{-1} \Upsilon(t) + \int_{t-\delta}^t \beta^{\frac{-\vartheta+t-\delta}{\delta}} b(\zeta(\vartheta)) d\vartheta \\ &\triangleq \beta^{-1} (\Upsilon(t) + b_c) \end{aligned} \quad (5.21)$$

where $b_c = \int_{t-\delta}^t \beta^{\frac{-\vartheta+t-\delta}{\delta}} b(\zeta(\vartheta)) d\vartheta$ is the cost value for $[t - \delta, t)$, $b_c = [b_{c_{N_l+1}}^T \ b_{c_{N_l+2}}^T \ \dots \ b_{c_{N_l+N_f}}^T]^T$ with

$$b_{c_i} = \int_{t-\delta}^t \beta^{\frac{-\zeta+t}{\delta}} b(\zeta_i(\vartheta)) d\vartheta \quad (5.22)$$

$$= \begin{cases} 0 & \text{if } \zeta_i(t)^2 \leq h_i \\ \frac{\delta}{\ln(\beta)}(\beta - 1) & \text{if } \zeta_i(t)^2 > h_i \end{cases} \quad (5.23)$$

Due to the fact that (5.19) encompasses the future system information, it is hard to get the solution at the current time. To do away with such difficulty, a critic RBFNN is deployed to approximate it introduced to approximate it.

$$\Upsilon(t) = V_c^T \Theta_c(Y_c(t)) + \epsilon_c(Y_c(t)) \quad (5.24)$$

The estimate of $\Upsilon(t)$ and $\hat{\Upsilon}(t - \delta)$ learnt by the critic are described by:

$$\hat{\Upsilon}(t) = \hat{V}_c^T \Theta_c(Y_c(t)) \quad (5.25)$$

$$\hat{\Upsilon}(t - \delta) = \hat{V}_c^T \Theta_c(Y_c(t - \delta)) \quad (5.26)$$

The temporal difference error is calculated as follows:

$$\begin{aligned} E_c &= \hat{\Upsilon}(t) - \beta \hat{\Upsilon}(t - \delta) + b_c \\ &= \hat{V}_c^T \Delta \Theta_c(t) + b_c \\ &= \tilde{V}_c^T \Delta \Theta_c(t) + b_c + V_c^T \Delta \Theta_c(t) \end{aligned} \quad (5.27)$$

where $\tilde{V}_c = \hat{V}_c - V_c$ is the estimation error, $\Delta \Theta_c(t) = [\Theta_c(Y_c(t)) - \beta \Theta_c(Y_c(t - \delta))]$.

Consider the following cost function:

$$I_c = \frac{1}{2} E_c^T E_c \quad (5.28)$$

To minimize (5.28), the gradient-decent technique is adopted which yields:

$$\hat{V}_c = -\alpha_c \Delta \Theta_c(t) \left[\hat{V}_c^T \Delta \Theta_c(t) + b_c \right]^T - \eta_c \alpha_c \hat{V}_c - \eta_c \alpha_c \hat{V}_c^\gamma \quad (5.29)$$

where $\alpha_c = \text{diag}\{\alpha_{c_{N_l+1}} \ \alpha_{c_{N_l+2}} \dots \alpha_{c_{N_l+N_f}}\} > 0$ is the matrix of the learning rates, $\eta_c > 0$ is a small parameter, $\gamma \in (0, 1)$. The terms $-\eta_c \alpha_c \hat{V}_c - \eta_c \alpha_c \hat{V}_c^\gamma$ are added to (5.29) to improve its robustness and eliminate the need to guarantee the persistence excitation condition [134]

5.3.2 Actor design

It is not possible to design a controller for the system (5.18) since the functions F and G are not available. Therefore, the actor neural networks will be utilized to approximate them as follows:

$$F = V_{a1}^T \Theta_{a1}(Y_{a1}) + \epsilon_{a1}(Y_{a1}) \quad (5.30)$$

$$G = V_{a2}^T \Theta_{a2}(Y_{a2}) + \epsilon_{a2}(Y_{a2}) \quad (5.31)$$

The estimation of F and G learned by the two actor neural networks are thus:

$$\hat{F} = \hat{V}_{a1}^T \Theta_{a1}(Y_{a1}) \quad (5.32)$$

$$\hat{G} = \hat{V}_{a2}^T \Theta_{a2}(Y_{a2}) \quad (5.33)$$

It is required that the actor network ensures that $\check{q}_f$ tracks q_f^* with acceptable accuracy, and minimize $\hat{\Upsilon}(t)$ to the target cost value $\Upsilon_d = 0$. The error in the actor network is expressed as:

$$E_{a1} = \zeta_f + \hat{\Upsilon}(t) - \Upsilon_d = \zeta_f + \hat{V}_c^T \Theta_c(Y_c) \quad (5.34)$$

The actor cost function is defined as follows:

$$I_{a1} = \frac{1}{2} E_{a1}^T E_{a1} \quad (5.35)$$

By using the gradient-decent approach, the weight update law that minimizes (5.35) is calculated as:

$$\dot{\hat{V}}_{a1} = \alpha_{a1} \Theta_{a1}(Y_{a1}) \left[\zeta_f + \hat{V}_c^T \Theta_c(Y_c) \right]^T - \eta_{a1} \alpha_{a1} \hat{V}_{a1} - \eta_{a1} \alpha_{a1} \hat{V}_{a1}^\gamma \quad (5.36)$$

where $\alpha_{a1} = \text{diag}\{\alpha_{a1N_l+1} \ \alpha_{a1N_l+2} \ \dots \ \alpha_{a1N_l+N_f}\} > 0$ is the matrix of learning rates, $\eta_{a1} > 0$ is a small parameter. The term $-\eta_{a1} \alpha_{a1} \hat{V}_{a1}$ is included in (5.36) to enhance its robustness.

Following the same procedure for designing (5.36), the weight update law for $\hat{V}_{a2}$ is established as:

$$\dot{\hat{V}}_{a2} = \alpha_{a2} \Theta_{a2}(Y_{a2}) \left[\zeta_f \tau + \hat{V}_c^T \Theta_c(Y_c) \right]^T - \eta_{a2} \alpha_{a2} \hat{V}_{a2} - \eta_{a2} \alpha_{a2} \hat{V}_{a2}^\gamma \quad (5.37)$$

where $\alpha_{a2} = \text{diag}\{\alpha_{a2N_l+1} \ \alpha_{a2N_l+2} \ \dots \ \alpha_{a2N_l+N_f}\} > 0$ is the matrix of the learning rates,

$\eta_{a2} > 0$ is a small parameter.

Equation (5.18) can be rewritten as follows:

$$\dot{\zeta}_f = \bar{M}_{ff} \left\{ F - \tilde{G}\tau + \hat{G}\tau - \ddot{q}_f^* \right\} + \Lambda \dot{z}_f + \rho \frac{b_2}{b_1} \text{diag}\{\dot{z}_f\} z_f^{\frac{b_2}{b_1}-1} \quad (5.38)$$

where $\tilde{G}\tau = \hat{G}\tau - G\tau$ is the neural network approximation error.

Evaluating (5.38) gives:

$$\begin{aligned} \bar{M}_{ff}^{-1} \dot{\zeta}_f &= V_{a1}^T \Theta_{a1}(Y_{a1}) - \tilde{V}_{a2}^T \Theta_{a2}(Y_{a2}) \tau + \hat{G}\tau - \ddot{q}_f^* \\ &+ \left\{ \epsilon_{a1}(Y_{a1}) + \epsilon_{a2}(Y_{a2}) \tau \right\} + \Lambda \bar{M}_{ff}^{-1} \dot{z}_f \\ &+ \rho \frac{b_2}{b_1} \text{diag}\{\dot{z}_f\} \bar{M}_{ff}^{-1} z_f^{\frac{b_2}{b_1}-1} \end{aligned} \quad (5.39)$$

where $\tilde{V}_{a2} = \hat{V}_{a2} - V_{a2}$ is the weight estimation error.

The equivalent control command is designed as follows:

$$\tau_{eq} = \frac{1}{\hat{G}} \begin{bmatrix} -\hat{V}_{a1}^T \Theta_{a1}(Y_{a1}) - \rho \frac{b_2}{b_1} \text{diag}\{\dot{z}_f\} \bar{M}_{ff}^{-1} z_f^{\frac{b_2}{b_1}-1} \\ -\Lambda \bar{M}_{ff}^{-1} \dot{z}_f + \ddot{q}_f^* \end{bmatrix} \quad (5.40)$$

The reaching control command is included in the equivalent control command to augment the leader-follower performance as follows:

$$\tau = \tau_{eq} - \frac{1}{\hat{G}} [K_1 \zeta_f + K_2 \text{sgn}^\gamma(\zeta_f)] \quad (5.41)$$

where $K_1, K_2 \in \mathbb{R}^{nN \times nN}$ are constant diagonal matrices, $\text{sgn}^\gamma(\zeta_f) = \text{diag}\{\text{sgn}^\gamma(\zeta_{N_l+1}) \text{sgn}^\gamma(\zeta_{N_l+2}) \dots \text{sgn}^\gamma(\zeta_{N_l+N_f})\}$, $\text{sgn}^\gamma(\zeta_f) \triangleq \text{sgn}(\zeta_f) |\zeta_f|^\gamma$, $\gamma \in (0, 1)$.

By substituting (5.41) into (5.39), we have:

$$\begin{aligned}\bar{M}_{ff}^{-1}\dot{\zeta}_f &= -\tilde{V}_{a1}^T\Theta_{a1}(Y_{a1}) - \tilde{V}_{a2}^T\Theta_{a2}(Y_{a2})\tau - K_1\zeta_f \\ &\quad - K_2 \operatorname{sgn}(\zeta_f)|\zeta_f|^\gamma + \left\{ \epsilon_{a1}(Y_{a1}) + \epsilon_{a2}(Y_{a2})\tau \right\}\end{aligned}\quad (5.42)$$

where $\tilde{V}_{a1} = \hat{V}_{a1} - V_{a1}$ is the weight estimation error $|\zeta_f| = \left[|\zeta_{N_l+1}|^T |\zeta_{N_l+2}|^T \dots |\zeta_{N_l+N_f}|^T \right]^T$

Theorem 5.1 *If Assumptions (1-4) are satisfied, given the critic neural network (5.24) and actor neural networks (5.30) and (5.31), then the control law (5.41) with neural network update laws (5.36), (5.31) and (5.29) guarantees that all the closed-loop error signals are semiglobally uniformly ultimately bounded and the leader-follower AFM can be achieved. Moreover, the closed-loop error signals are trapped within a compact set defined as follows:*

$$\|\zeta_f\| \leq \min \left\{ \sqrt{2 \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{2 \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.43)$$

$$\|\tilde{V}_{a1}\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_{a1})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_{a1})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.44)$$

$$\|\tilde{V}_{a2}\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_{a2})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_{a2})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.45)$$

$$\|\tilde{V}_c\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_c)^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_c)^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.46)$$

Proof. Select the following Lyapunov candidate function:

$$\begin{aligned}
L &= \frac{1}{2} \zeta_f^T \bar{M}_{ff}^{-1} \zeta_f + \frac{1}{2} \left(\tilde{V}_{a1}^T \alpha_{a1}^{-1} \tilde{V}_{a1} \right) \\
&\quad + \frac{1}{2} \left(\tilde{V}_{a2}^T \alpha_{a2}^{-1} \tilde{V}_{a2} \right) + \frac{\Lambda_c}{2} \left(\tilde{V}_c^T \alpha_c^{-1} \tilde{V}_c \right) \\
&\triangleq L_1 + L_2 + L_3
\end{aligned} \tag{5.47}$$

where

$$\begin{aligned}
L_1 &= \frac{1}{2} \zeta_f^T \bar{M}_{ff}^{-1} \zeta_f \\
L_2 &= \frac{1}{2} \left(\tilde{V}_{a1}^T \alpha_{a1}^{-1} \tilde{V}_{a1} \right) + \frac{1}{2} \left(\tilde{V}_{a2}^T \alpha_{a2}^{-1} \tilde{V}_{a2} \right) \\
L_3 &= \frac{\Lambda_c}{2} \left(\tilde{V}_c^T \alpha_c^{-1} \tilde{V}_c \right)
\end{aligned}$$

One can obtain that:

$$\begin{aligned}
\dot{L}_1 &= \zeta_f^T \dot{\bar{M}}_{ff}^{-1} \zeta_f \\
&= -\zeta_f^T K_1 \zeta_f - \zeta_f^T K_2 \operatorname{sgn}(\zeta_f) |\zeta_f|^\gamma - \zeta_f^T \tilde{V}_{a1}^T \Theta_{a1}(Y_{a1}) \\
&\quad - \zeta_f^T \tilde{V}_{a2}^T \Delta \Theta_{a2}(Y_{a2}) \tau + \zeta_f^T \epsilon_{a1}(Y_{a1}) + \zeta_f^T \epsilon_{a2}(Y_{a2}) \tau
\end{aligned} \tag{5.48}$$

$$\begin{aligned}
\dot{L}_2 &= \left(\tilde{V}_{a1}^T \alpha_{a1}^{-1} \dot{\tilde{V}}_{a1} \right) + \left(\tilde{V}_{a2}^T \alpha_{a2}^{-1} \dot{\tilde{V}}_{a2} \right) \\
&= \left(\tilde{V}_{a1}^T \alpha_{a1}^{-1} \hat{\dot{V}}_{a1} \right) + \left(\tilde{V}_{a2}^T \alpha_{a2}^{-1} \hat{\dot{V}}_{a2} \right) \\
&= -\eta_{a1} \tilde{V}_{a1}^T \hat{V}_{a1} - \eta_{a2} \tilde{V}_{a2}^T \hat{V}_{a2} - \eta_{a1} \tilde{V}_{a1}^T \hat{V}_{a1}^\gamma - \eta_{a2} \tilde{V}_{a2}^T \hat{V}_{a2}^\gamma \\
&\quad + \zeta_f^T \tilde{V}_{a1}^T \Delta \Theta_{a1}(Y_{a1}) + \zeta_f^T \tilde{V}_{a2}^T \Delta \Theta_{a2}(Y_{a2}) \\
&\quad + \tilde{V}_{a1}^T \Delta \Theta_{a1}(Y_{a1}) [\hat{V}_c^T \Delta \Theta_c(Y_c)]^T \\
&\quad + \tilde{V}_{a2}^T \Delta \Theta_{a2}(Y_{a2}) [\hat{V}_c^T \Delta \Theta_c(Y_c)]^T
\end{aligned} \tag{5.49}$$

$$\begin{aligned}
\dot{L}_1 + \dot{L}_2 = & -\zeta_f^T K_1 \zeta_f - \zeta_f^T K_2 \operatorname{sgn}(\zeta_f) |\zeta_f|^\gamma - \zeta_f^T \tilde{V}_{a1}^T \Theta_{a1}(Y_{a1}) \\
& - \zeta_f^T \tilde{V}_{a2}^T \Delta \Theta_{a2}(Y_{a2}) \tau + \zeta_f^T \epsilon_{a1} + \zeta_f^T \epsilon_{a2} \\
& - \eta_{a1} \tilde{V}_{a1}^T [\tilde{V}_{a1} + V_{a1}] - \eta_{a2} \tilde{V}_{a2}^T [\tilde{V}_{a2} + V_{a2}] \\
& - \eta_{a1} \tilde{V}_{a1}^T [\tilde{V}_{a1} + V_{a1}]^\gamma - \eta_{a2} \tilde{V}_{a2}^T [\tilde{V}_{a2} + V_{a2}]^\gamma \\
& + \tilde{V}_{a1}^T \Delta \Theta_{a1}(Y_{a1}) [[\tilde{V}_c + V_c]^T \Delta \Theta_c(Y_c)]^T \\
& + \tilde{V}_{a2}^T \Delta \Theta_{a2}(Y_{a2}) [[\tilde{V}_c + V_c]^T \Delta \Theta_c(Y_c)]^T
\end{aligned} \tag{5.50}$$

Considering Lemma 5.1, one has:

$$-\tilde{V}_{a1}^T [\tilde{V}_{a1} + V_{a1}]^\gamma \leq -\frac{1-\varrho}{1+\gamma} \tilde{V}_{a1}^{1+\gamma} + \frac{\bar{V}_{a1}}{1+\gamma} \tag{5.51}$$

$$-\eta_{a2} \tilde{V}_{a2}^T [\tilde{V}_{a2} + V_{a2}]^\gamma \leq -\frac{1-\varrho}{1+\gamma} \tilde{V}_{a2}^{1+\gamma} + \frac{\bar{V}_{a2}}{1+\gamma} \tag{5.52}$$

where

$$\begin{aligned}
\bar{V}_{a1} &= V_{a1}^{1+\gamma} + \left(\frac{V_{a1}}{1 - (1-\varrho)^{1/(1+\gamma)}} \right)^{1+\gamma} + \left(\frac{V_{a1}(1-\varrho)^{1/(1+\gamma)}}{1 - (1-\varrho)^{1/(1+\gamma)}} \right)^{1+\gamma} \\
\bar{V}_{a2} &= V_{a2}^{1+\gamma} + \left(\frac{V_{a2}}{1 - (1-\varrho)^{1/(1+\gamma)}} \right)^{1+\gamma} + \left(\frac{V_{a2}(1-\varrho)^{1/(1+\gamma)}}{1 - (1-\varrho)^{1/(1+\gamma)}} \right)^{1+\gamma}
\end{aligned}$$

Substituting the inequalities (5.51) and (5.52) into (5.50) gives:

$$\begin{aligned}
\dot{L}_1 + \dot{L}_2 \leq & -\left(\lambda_{\min}(K_1) - \frac{\kappa_1 + \kappa_2}{2} - 2\right)\|\zeta_f\|^2 - \lambda_{\min}(K_2)|\zeta_f|^{\gamma+1} \\
& - \left(\frac{\eta_{a1}}{2} - \frac{\kappa_1}{2} - \kappa_1\kappa_2\right)\|\tilde{V}_{a1}\|^2 - \eta_{a1}\frac{1-\varrho}{1+\gamma}\left(\|\tilde{V}_{a1}\|^2\right)^{\frac{1+\gamma}{2}} \\
& - \left(\frac{\eta_{a2}}{2} - \frac{\kappa_2}{2} - \kappa_1\kappa_2\right)\|\tilde{V}_{a2}\|^2 - \eta_{a2}\frac{1-\varrho}{1+\gamma}\left(\|\tilde{V}_{a2}\|^2\right)^{\frac{1+\gamma}{2}} \\
& + \frac{\kappa_1\kappa_c + \kappa_2\kappa_c}{2}\|\tilde{V}_c\|^2 + \mathcal{Q}_a
\end{aligned} \tag{5.53}$$

where $\|\cdot\|$ is the Frobenius norm, $\|\Delta\Theta_{a1}\| \leq \kappa_1$, $\|\Delta\Theta_{a2}\| \leq \kappa_2$, $\|\Delta\Theta_c\| \leq \kappa_c$, $\|V_{a1}\| \leq v_{a1}$, $\|V_{a2}\| \leq v_{a2}$, $\|V_c\| \leq v_c$, $\epsilon_{a1} \leq \bar{\epsilon}_{a1}$, $\epsilon_{a2} \leq \bar{\epsilon}_{a2}$, $\mathcal{Q}_a = \frac{\bar{V}_{a1}}{1+\gamma} + \frac{\bar{V}_{a2}}{1+\gamma} + \frac{\bar{\epsilon}_{a1}^2}{2} + \frac{\bar{\epsilon}_{a2}^2}{2} + \frac{\kappa_1\kappa_c + \kappa_2\kappa_c}{2}v_c^2 + \frac{\eta_{a1}}{2}v_{a1}^2 + \frac{\eta_{a2}}{2}v_{a2}^2$, with v_{a1} , v_{a2} , v_c , κ_1 , κ_2 , κ_c , $\bar{\epsilon}_{a1}$, and $\bar{\epsilon}_{a2}$ being bound constants.

The time-derivative of L_3 yields:

$$\begin{aligned}
\dot{L}_3 = & \Lambda_c \left(\tilde{V}_c^T \alpha_c^{-1} \dot{\tilde{V}}_c \right) = \Lambda_c \left(\tilde{V}_c^T \alpha_c^{-1} \hat{\dot{V}}_c \right) \\
= & -\Lambda_c \tilde{V}_c^T \Delta\Theta_c(t) \left[\hat{V}_c^T \Delta\Theta_c(t) + b_c \right]^T \\
& - \eta_c \Lambda_c \tilde{V}_c^T \hat{V}_c - \eta_c \Lambda_c \tilde{V}_c^T \hat{V}_c^\gamma \\
= & -\Lambda_c \tilde{V}_c^T \Delta\Theta_c(t) \left[\tilde{V}_c^T \Delta\Theta_c(t) + V_c^T \Delta\Theta_c(t) + b_c \right]^T \\
& - \eta_c \Lambda_c \tilde{V}_c^T \tilde{V}_c - \eta_c \Lambda_c \tilde{V}_c^T V_c - \eta_c \Lambda_c \tilde{V}_c^T [\tilde{V}_c + V_c]^\gamma
\end{aligned} \tag{5.54}$$

Based on Lemma 5.1, we get:

$$-\tilde{V}_c^T [\tilde{V}_c + V_c]^\gamma \leq -\frac{1-\varrho}{1+\gamma} \tilde{V}_c^{1+\gamma} + \frac{\bar{V}_c}{1+\gamma} \tag{5.55}$$

where

$$\bar{V}_c = V_c^{1+\gamma} + \left(\frac{V_c}{1 - (1 - \varrho)^{1/(1+\gamma)}} \right)^{1+\gamma} + \left(\frac{V_c(1 - \varrho)^{1/(1+\gamma)}}{1 - (1 - \varrho)^{1/(1+\gamma)}} \right)^{1+\gamma}$$

Combining (5.54) with (5.55) yields:

$$\begin{aligned} \dot{L}_3 &\leq -\Lambda_c \left\{ \eta_c + \|\Delta\Theta_c(t)\|^2 \right\} \|\tilde{V}_c\|^2 - \left\{ \|V_c\| \|\Delta\Theta_c(t)\|^2 \right. \\ &\quad \left. + \|\Delta\Theta_c(t)\| \|b_c\| + \eta_c \|V_c\| \right\} \Lambda_c \|\tilde{V}_c\| - \eta_c \Lambda_c \left(\|\tilde{V}_c\|^2 \right)^{\frac{1+\gamma}{2}} + \bar{V}_c \\ &\leq -\eta_c \Lambda_c \|\tilde{V}_c\|^2 + \mathcal{Q}_c \|\tilde{V}_c\| - \eta_c \Lambda_c \left(\|\tilde{V}_c\|^2 \right)^{\frac{1+\gamma}{2}} + \bar{V}_c \\ &\leq -\left\{ \eta_c \Lambda_c + \Lambda_c \|\Delta\Theta_c(t)\|^2 - \frac{1}{2} \right\} \|\tilde{V}_c\|^2 \\ &\quad - \eta_c \Lambda_c \frac{1 - \varrho}{1 + \gamma} \left(\|\tilde{V}_c\|^2 \right)^{\frac{1+\gamma}{2}} + \frac{\bar{V}_c}{1 + \gamma} + \frac{\mathcal{Q}_c^2}{2} \end{aligned} \tag{5.56}$$

where $\mathcal{Q}_c = -\Lambda_c(v_c\kappa_c^2 + \kappa_c\kappa_{b_c} + \eta_cv_c)$, with $\|b_c\| \leq \kappa_{b_c}$.

Finally, the time-derivative of L is achieved as follows:

$$\begin{aligned} \dot{L} &\leq -\left(\lambda_{\min}(K_1) - \frac{\kappa_1 + \kappa_2}{2} - 2 \right) \|\zeta_f\|^2 - \lambda_{\min}(K_2) |\zeta_f|^{\gamma+1} \\ &\quad - \left(\frac{\eta_{a1}}{2} - \frac{\kappa_1}{2} - \kappa_1\kappa_2 \right) \|\tilde{V}_{a1}\|^2 - \eta_{a1} \frac{1 - \varrho}{1 + \gamma} \left(\|\tilde{V}_{a1}\|^2 \right)^{\frac{1+\gamma}{2}} \\ &\quad - \left(\frac{\eta_{a2}}{2} - \frac{\kappa_2}{2} - \kappa_1\kappa_2 \right) \|\tilde{V}_{a2}\|^2 - \eta_{a2} \frac{1 - \varrho}{1 + \gamma} \left(\|\tilde{V}_{a2}\|^2 \right)^{\frac{1+\gamma}{2}} \\ &\quad - \left(\Lambda_c\eta_c + \Lambda_c\kappa_c^2 - \frac{1}{2} - \frac{\kappa_1\kappa_c + \kappa_2\kappa_c}{2} \right) \|\tilde{V}_c\|^2 \\ &\quad - \eta_c \Lambda_c \frac{1 - \varrho}{1 + \gamma} \left(\|\tilde{V}_c\|^2 \right)^{\frac{1+\gamma}{2}} + \frac{\bar{V}_c}{1 + \gamma} + \frac{\mathcal{Q}_c^2}{2} + \mathcal{Q}_a \end{aligned} \tag{5.57}$$

The parameters $\Lambda_c > 0$, $\Lambda > 0$, $\eta_{a1} > 0$, $\eta_{a2} > 0$, $\eta_c > 0$, and the matrices $K_1 > 0$,

$K_2 > 0$ are chosen such that the following inequalities are valid:

$$\begin{aligned}
& \lambda_{\min}(K_1) > 0; \quad \lambda_{\min}(K_2) > 0; \\
& \left(\lambda_{\min}(K_1) - \frac{\kappa_1 + \kappa_2}{2} - 2 \right) > 0 \\
& \left(\frac{\eta_{a1}}{2} - \frac{\kappa_1}{2} - \kappa_1 \kappa_2 \right) > 0 \\
& \left(\frac{\eta_{a2}}{2} - \frac{\kappa_2}{2} - \kappa_1 \kappa_2 \right) > 0 \\
& \left(\Lambda_c \eta_c + \Lambda_c \kappa_c^2 - \frac{1}{2} - \frac{\kappa_1 \kappa_c + \kappa_2 \kappa_c}{2} \right) > 0
\end{aligned}$$

Equation (5.57) can be simplified as:

$$\dot{L} \leq -Q_1 L - Q_2 L^{\frac{\gamma+1}{2}} + Q_3 \quad (5.58)$$

where

$$\begin{aligned}
\mathcal{Q}_1 &= \min \left\{ 2 \left(\lambda_{\min}(K_1) - \frac{\kappa_1 + \kappa_2}{2} - 2 \right), \right. \\
& \quad 2 \left(\frac{\eta_{a1}}{2} - \frac{\kappa_1}{2} - \kappa_1 \kappa_2 \right), 2 \left(\frac{\eta_{a2}}{2} - \frac{\kappa_2}{2} - \kappa_1 \kappa_2 \right), \\
& \quad \left. 2 \left(\Lambda_c \eta_c + \Lambda_c \kappa_c^2 - \frac{1}{2} - \frac{\kappa_1 \kappa_c + \kappa_2 \kappa_c}{2} \right) \right\} \\
\mathcal{Q}_2 &= \min \left\{ 2 \lambda_{\min}(K_2), 2 \alpha_{a1} \eta_c \frac{1 - \varrho}{1 + \gamma}, 2 \alpha_{a2} \eta_c \frac{1 - \varrho}{1 + \gamma}, \right. \\
& \quad \left. 2 \alpha_c \eta_c \Lambda_c \frac{1 - \varrho}{1 + \gamma} \right\} \\
\mathcal{Q}_3 &= \frac{\bar{V}_2}{1 + \gamma} + \frac{\mathcal{Q}_c^2}{2} + \mathcal{Q}_a
\end{aligned}$$

Further one can achieve:

$$\dot{L} \leq -\mu \mathcal{Q}_1 L - (1 - \mu) \mathcal{Q}_1 L - \mathcal{Q}_2 L^{\frac{\gamma+1}{2}} + \mathcal{Q}_3 \quad (5.59)$$

or

$$\dot{L} \leq -\mathcal{Q}_1 L - (1 - \mu) \mathcal{Q}_2 L^{\frac{\gamma+1}{2}} - \mu \mathcal{Q}_2 L^{\frac{\gamma+1}{2}} + \mathcal{Q}_3 \quad (5.60)$$

where $0 < \mu < 1$. If $L > \frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1}$, then (5.59) yields:

$$\dot{L} \leq -\mu \mathcal{Q}_1 L - \mathcal{Q}_2 L^{\frac{\gamma+1}{2}} \quad (5.61)$$

Similarly, if $L^{\frac{\gamma+1}{2}} > \frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2}$, then (5.60) gives:

$$\dot{L} \leq -\mathcal{Q}_1 L - \mu \mathcal{Q}_2 L^{\frac{\gamma+1}{2}} \quad (5.62)$$

By using Lemma 5.2, the Lyapunov function L will decrease in finite time and converge into the following region:

$$L \leq \min \left\{ \frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1}, \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}} \right\} \quad (5.63)$$

Therefore, within the finite settling time T_s , all the error signals in the closed-loop system converge to the following regions:

$$\|\zeta_f\| \leq \min \left\{ \sqrt{2 \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{2 \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.64)$$

$$\|\tilde{V}_{a1}\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_{a1})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_{a1})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.65)$$

$$\|\tilde{V}_{a2}\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_{a2})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_{a2})^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.66)$$

$$\|\tilde{V}_c\| \leq \min \left\{ \sqrt{\frac{2}{\max(\lambda_c)^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_1} \right)}, \sqrt{\frac{2}{\max(\lambda_c)^{-1}} \left(\frac{\mathcal{Q}_3}{(1-\mu)\mathcal{Q}_2} \right)^{\frac{2}{1+\gamma}}} \right\} \quad (5.67)$$

■

Remark 5.2 Based on the definition of $\mathcal{Q}_1$, $\mathcal{Q}_2$ and $\mathcal{Q}_3$, the regions of convergence (5.64)-(5.67) can be made sufficiently small by selecting appropriate controller parameters Λ_c , Λ , η_{a1} , η_{a2} , $\eta_c > 0$, and the matrices K_1 and K_2 .

5.4 Simulation Results

This section provides a simulation example to validate the performance of the Secure Formation Maneuver Control protocol developed. The nonlinear dynamic model of an i th agent is given by:

$$\ddot{q}_i + C_i(q_i, \dot{q}_i) + H_i(q_i) = B_i(\tau_i) + \tau_{di} \quad (5.68)$$

where

$$C_i(q_i, \dot{q}_i) = \begin{bmatrix} q_{i1}^2 + 2\dot{q}_{i1} \cos(q_{i1})\dot{q}_{i1} \\ q_{i2}^2 + 2\dot{q}_{i2} \cos(q_{i2})\dot{q}_{i2} \end{bmatrix}$$

$$H_i(q_i) = \begin{bmatrix} 1 + 0.5\cos(\dot{q}_{i1}) & 0 \\ 0 & 1 + 0.5\cos(\dot{q}_{i2}) \end{bmatrix}$$

$$\tau_{di} = \begin{bmatrix} \sin(t) \\ \cos(t) \end{bmatrix}$$

The sensor deception attack is given by:

$$\check{q}_i = q_i + \varpi_i(t)q_i = \pi_i(t)q_i \quad (5.69)$$

where

$$\pi_i(t) = 1 + \varpi_i(t) = 1 + \begin{bmatrix} 2\cos(2t \ q_{i1}) & 0 \\ 0 & 3\cos(2t \ q_{i2}) \end{bmatrix}$$

The MAS under this study consists of three leaders and four followers. The nominal (target) formation of the agents is shown in Fig. 5.1. It consists of seven agents of which the agents marked 1,2 and 3 are the leaders while the remaining are the followers. From Fig. 5.1, it can be seen that the nominal configuration of the seven agents is $v_1 = [6 \ 0]$, $v_2 = [3 \ 3]$, $v_3 = [3 \ -3]$, $v_4 = [0 \ 3]$, $v_5 = [0 \ -3]$, $v_6 = [-3 \ 3]$ and $v_7 = [-3 \ -3]$. Adopting the approach discussed in [49], the stresses matrix is calculated as:

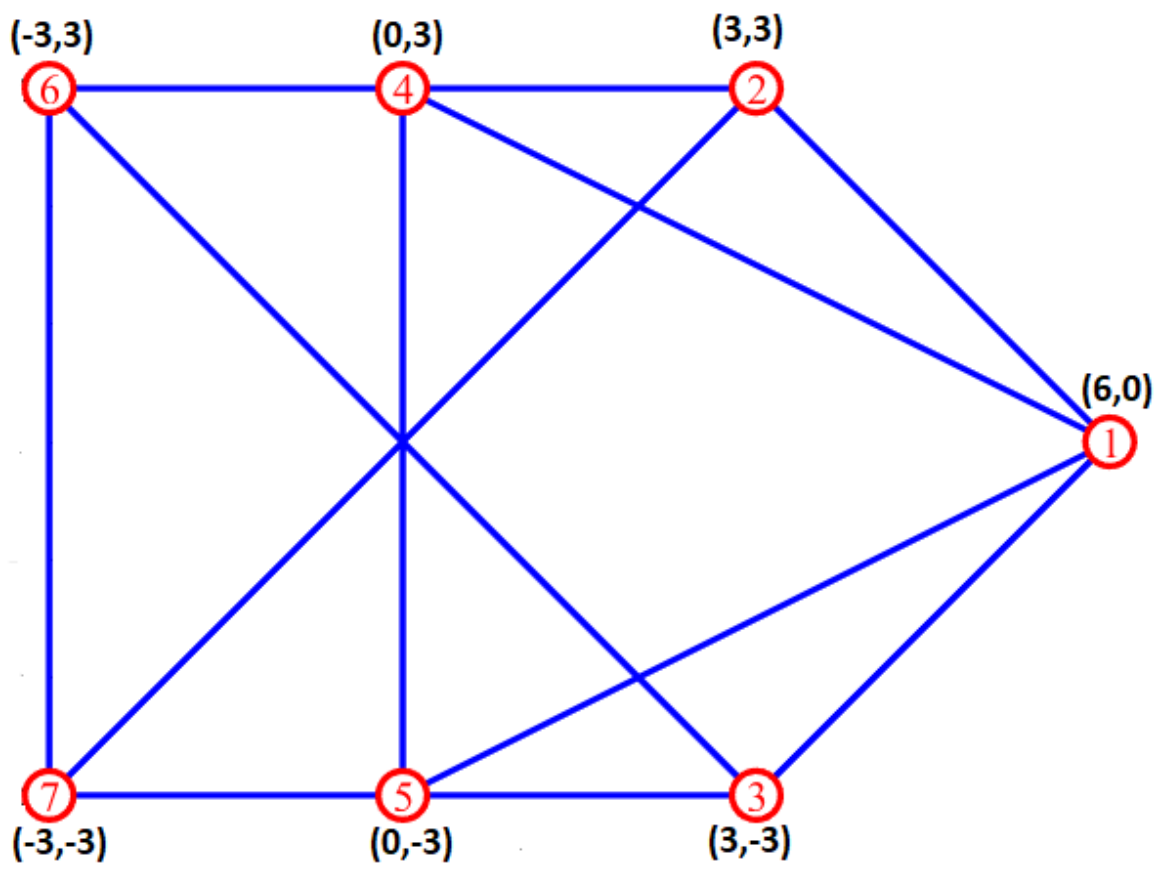


Figure 5.1: The nominal formation graph of the agents.

$$\Omega = \begin{bmatrix} -0.2741 & 0.2741 & 0.2741 & -0.1370 & -0.1370 & 0 & 0 \\ 0.2741 & -0.6852 & 0 & 0.5482 & 0 & 0 & -0.1370 \\ 0.2741 & 0 & -0.6852 & 0 & 0.5482 & -0.1370 & 0 \\ -0.1370 & 0.5482 & 0 & -0.7537 & 0.0685 & 0.2741 & 0 \\ -0.1370 & 0 & 0.5482 & 0.0685 & -0.7537 & 0 & 0.2741 \\ 0 & 0 & -0.1370 & 0.2741 & 0 & -0.2741 & 0.1370 \\ 0 & -0.1370 & 0 & 0 & 0.2741 & 0.1370 & -0.2741 \end{bmatrix} \quad (5.70)$$

The initial positions of the seven agents are given by $q_1(0) = [6 \ 0]^T$, $q_2(0) = [3 \ 3]^T$, $q_3(0) = [3 \ -3]^T$, $q_4(0) = [1 \ 4]^T$, $q_5(0) = [0 \ -2]^T$, $q_6(0) = [-4 \ 4]^T$, $q_7(0) = [-5 \ -4]^T$. The integral reinforcement interval is selected as $\tau = 0.1$, The controller gains are selected as, $\Lambda = 20$, $\rho = 2$, $b_1 = 0.8$, $b_2 = 1.44$, $\gamma = 0.5$, $K_1 = 25I$, and $K_2 = 8I$. The actor neural network parameters are chosen as $\eta_{a1} = \eta_{a2} = 0.2$, $\alpha_{a1} = \alpha_{a2} = 0.001$, the center ϖ_{a1} and ϖ_{a2} are selected between $1, 0.5, \dots, 5$, the width of the Gaussian basis function $\eta_{a1} = \eta_{a1} = 1$, number of nodes in the hidden layer is set as 5, the initial weights vector of the actor networks are chosen as $W_{a10} = [0.1, 0.1, \dots, 0.1]^T$ and $W_{a20} = [0.1, 0.1, \dots, 0.1]^T$. The critic neural network parameters are chosen as $\eta_c = 0.5$, $\alpha_c = 0.001$, the width of the Gaussian basis function $\eta_c = 1$, the center of the receptive field c_c is chosen between $1, 0.5, \dots, 5$, number of nodes in the hidden layer is as 5, the initial weight vector of the critic network is selected as $W_{c0} = [0.1, 0.1, \dots, 0.1]^T$.

The simulation results are presented in Fig. 5.2-5.7. The reinforcement learning-based finite-time fast terminal sliding mode control protocols capable of mitigating the sensor deception attacks and the actuator backlash-like hysteresis are depicted in Fig. 5.2. The tracking errors of the leader-follower MAS are depicted in Fig. 5.3. The

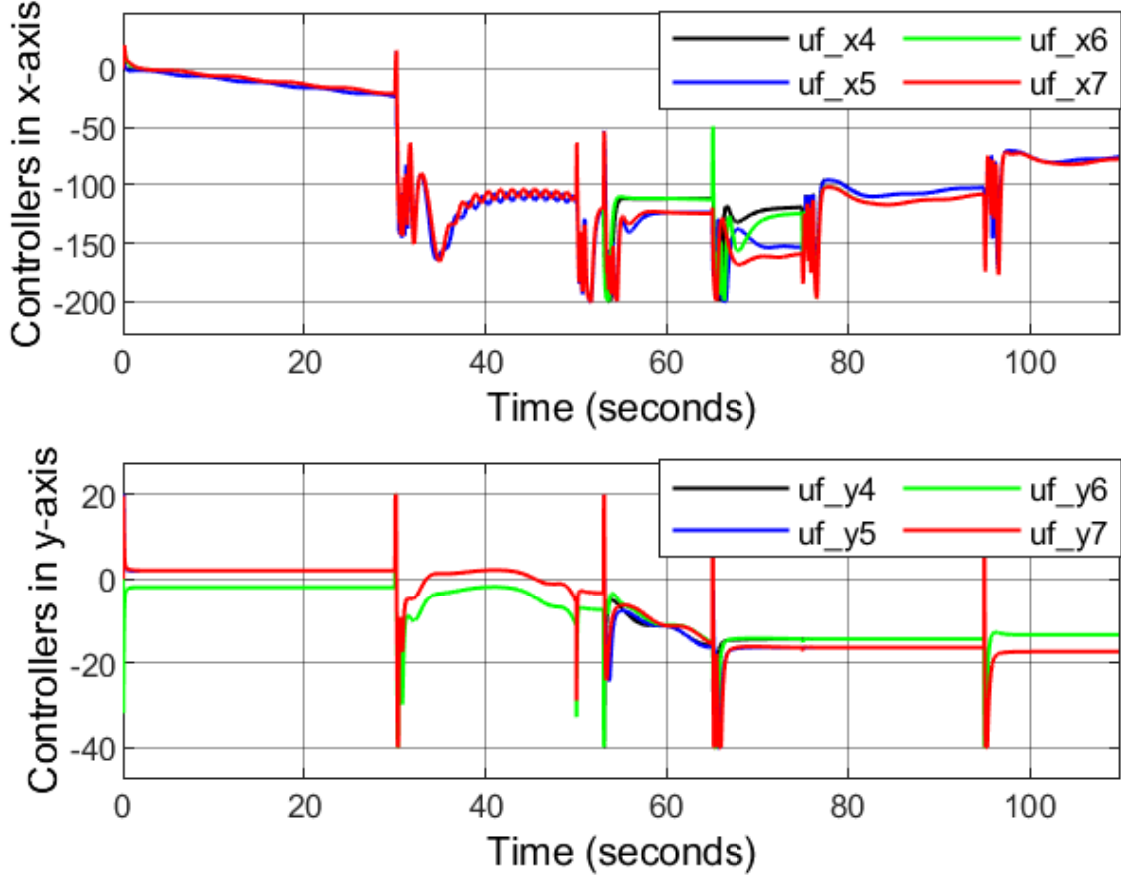


Figure 5.2: The finite-time secure control protocols

evolution of the norm-2 of the weight vectors of the actor neural network for estimating F and G are presented in Fig. 5.4 and Fig. 5.5, respectively. The evolution of the norm-2 of the weight vectors of the critic neural network is shown in Fig. 5.6. The various leader-follower AFMs of the MAS under deception attacks and backlash-like hysteresis are shown in Fig. 5.7. It can be seen that the adversaries could not degrade the formation of the agents due to the effectiveness of the proposed control strategy.

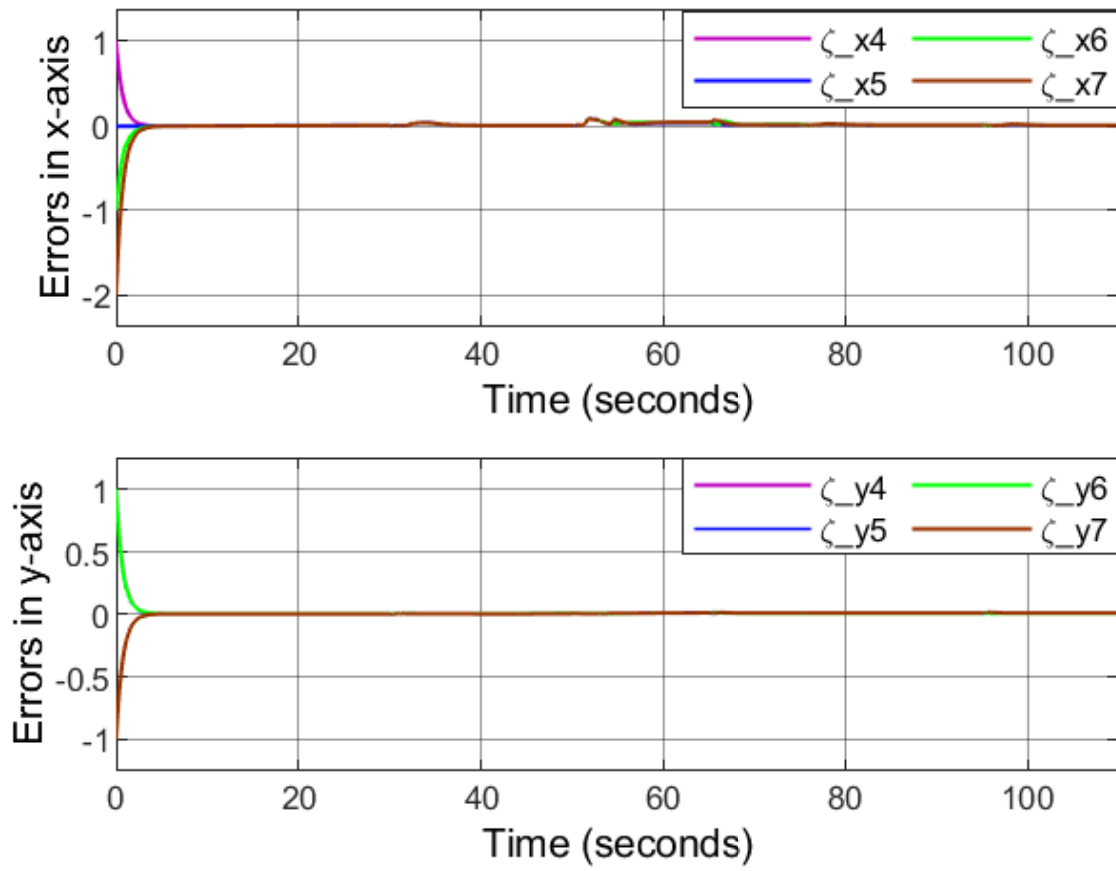


Figure 5.3: The tracking errors of the followers.

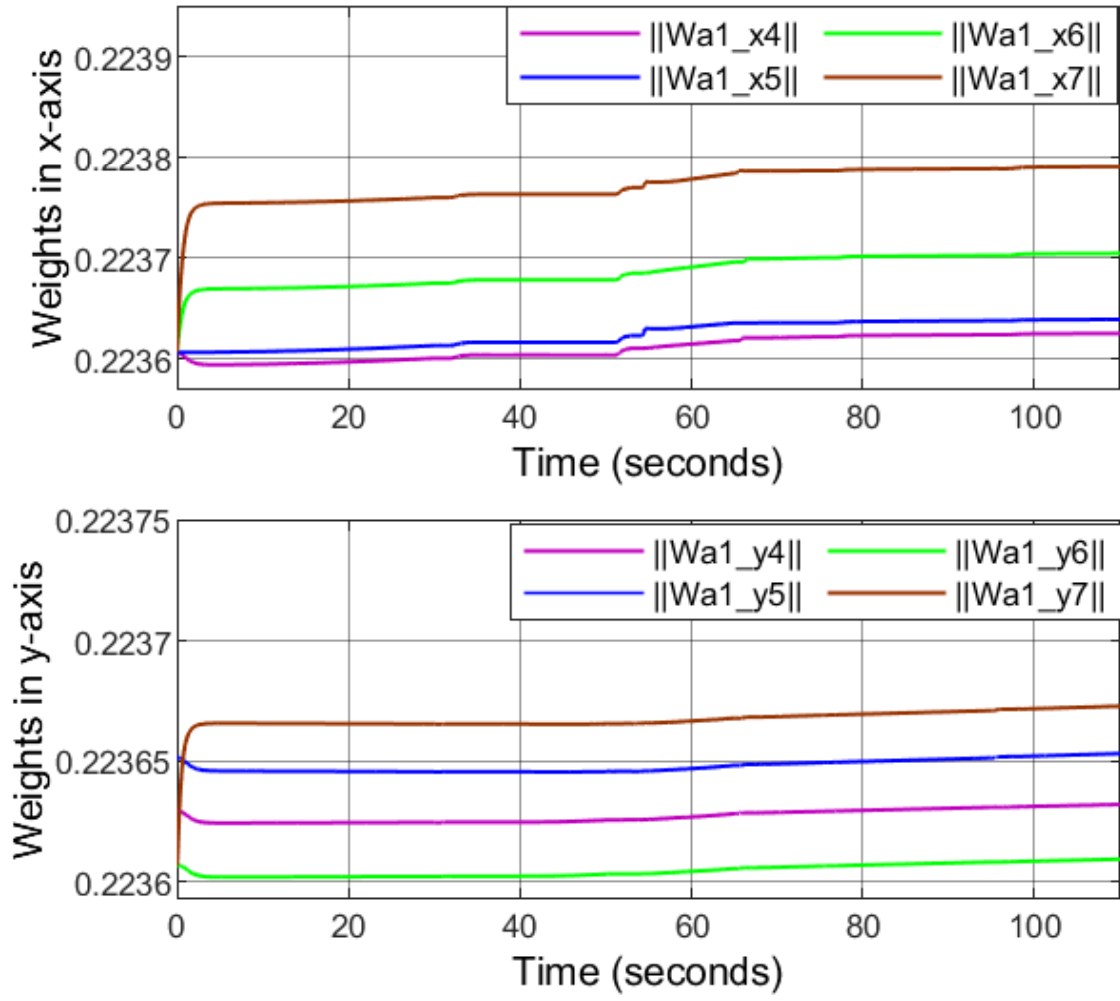


Figure 5.4: The norm-2 of weight vector of the actor neural network for estimating F .

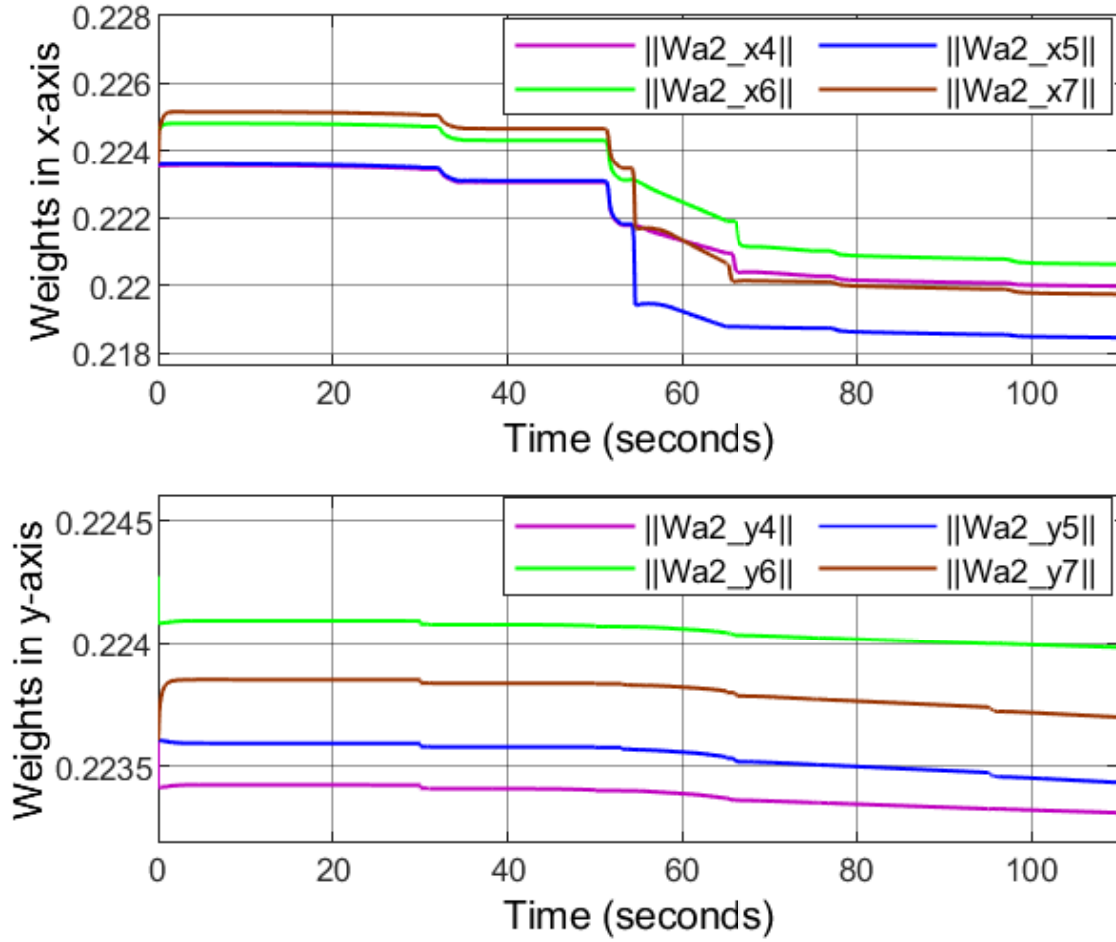


Figure 5.5: The norm-2 of weight vector of the actor neural network for estimating G .

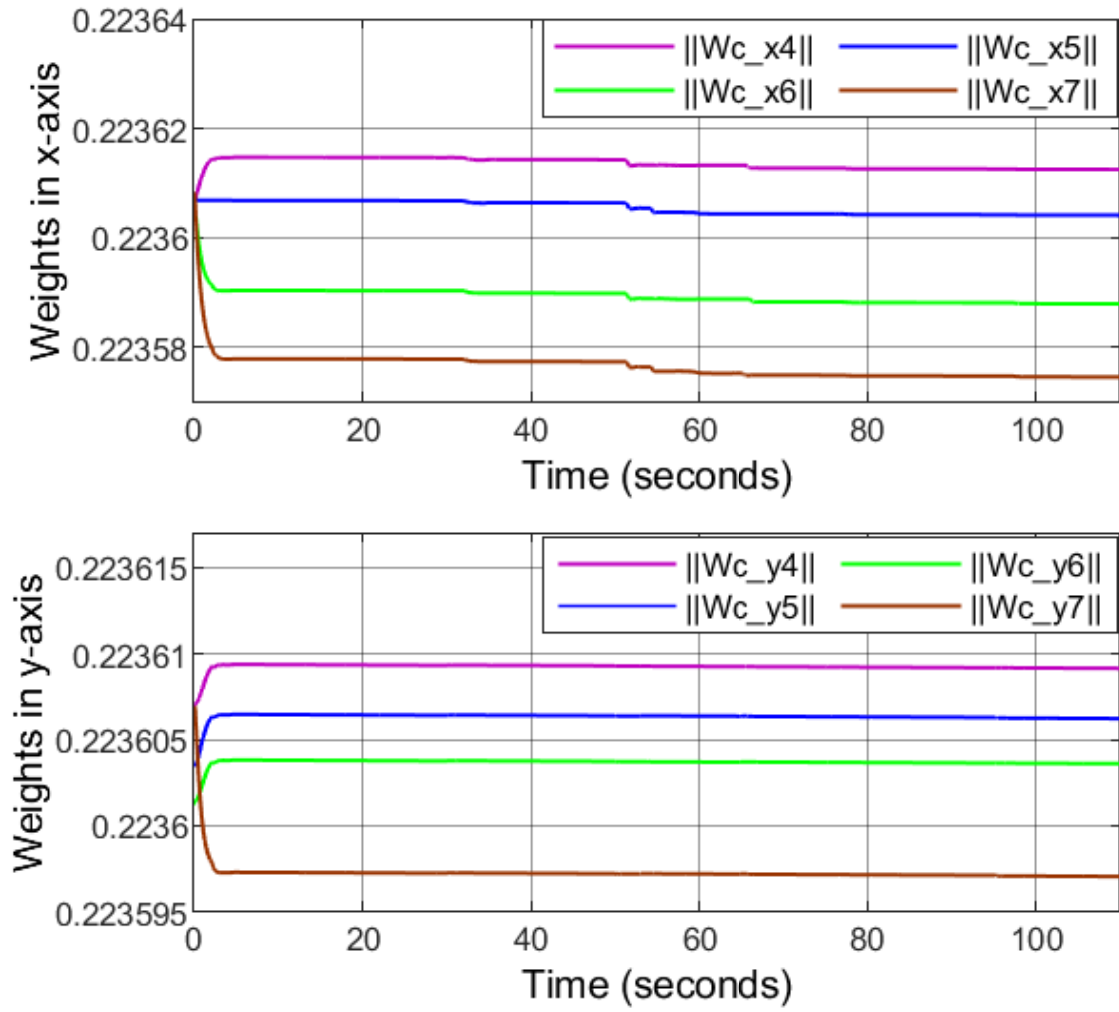


Figure 5.6: The norm-2 of weight vector of the critic neural network.

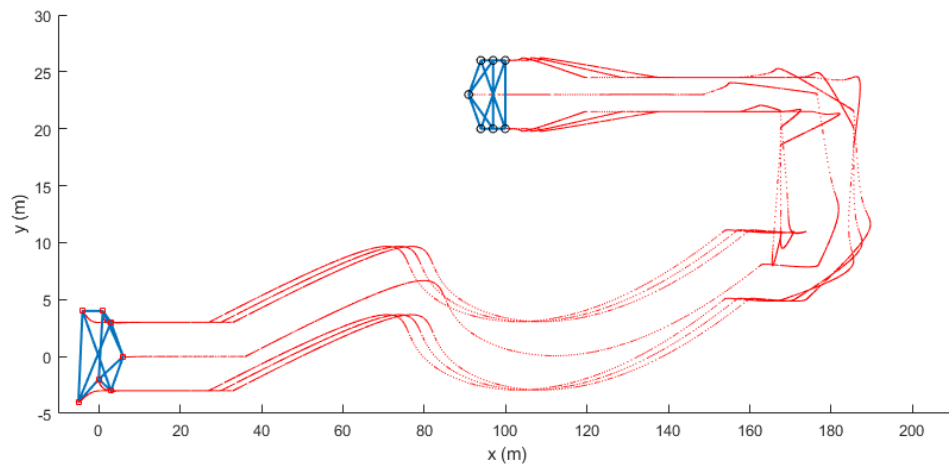


Figure 5.7: Various AFMs of the leader-follower MAS

CHAPTER 6

SECURE FAULT-TOLERANT REINFORCEMENT LEARNING-BASED PRESCRIBED PERFORMANCE AFFINE FORMATION MANEUVER CONTROL OF MULTI-AGENT SYSTEMS WITH SENSOR DECEPTION ATTACK

This chapter investigates the design of a novel prescribed-performance controller for fault-tolerant secure optimized backstepping control using reinforcement learning for affine formation maneuvering of second-order nonlinear multi-agent systems under sensor deception attacks and actuator faults. The control objective is to ensure that the leader-follower maneuvers are resilient in the event of deception attacks and actuator faults. Linear transformation matrices are used to obtain various formation maneuvers such as translation, rotation, scaling, shearing, and collinearity simultaneously. Whenever the desired formation maneuver changes, the dynamic motion of the agents undergo new transient states before converging to the new formation maneuver. In order to constrain the transient states of the formation tracking errors to a predefined small set whenever the desired formation maneuver changes, a modified prescribed performance function with finite settling time is introduced. The optimal backstepping control laws are obtained from the approximate solution of the

Hamilton-Jacobi-Bellman equation using reinforcement learning under the identifier-actor-critic neural networks. The resulting control algorithm is simple and does not require complex computations. Lyapunov candidate functions have been used to prove the boundedness of the closed-loop system. Simulation results have shown that the proposed controller can achieve various leader-follower affine formation maneuvers regardless of the sensor deception attacks and actuator faults.

6.1 Introduction

The utilization of wireless communication networks amongst agents offers multi-agent systems a particular edge in terms of simplicity, design cost, and efficiency [135]. Nevertheless, these advantages come at the expense of exposing the cyber-physical system to harmful cyber-attacks such as deception attacks [136], denial of service attacks [137], and random attacks [138]. The malicious signals injected by adversaries into the sensors and actuators of cyber-physical systems can deteriorate the desired formation pattern of the multi-agent systems and eventually destabilize them [139]. Therefore, secure and reliable control of multi-agent systems to track the desired formation geometry requires urgent attention. As of now, a series of results have been presented on the formation control of multi-agent systems subjected to denial of service attacks [140], [27], [141], [142], [143] but the dangers of deception attacks on the agent's formation received little attention [144]. Thus, this chapter considers affine formation maneuver control of multi-agent systems under deception attacks.

In addition to cyber-attacks, faults in the actuators of mobile agents can also deteriorate control performance or even lead to instability of the system [145]. Ad-

ditionally, in the context of the formation acquisition control of multi-agent systems, the actuator faults in one agent can spread over nearby agents via the communication network and impair the functionality of the entire system. Consequently, efforts have been made to address this problem through fault-tolerant control such that the formation acquired by the agents is not affected by faults [146], [147]. Yet, no prior research has addressed fault-tolerant control in the framework of affine formation maneuvering of multi-agent systems.

In practice, the formation control systems are required to meet certain performance specifications in transient and steady states [18]. The prescribed performance functions have been successfully employed to achieve these performance requirements by constraining the formation tracking errors to converge within a predefined small set. Various formation control problems of multi-agent systems with prescribed tracking error performances have been reported in the literature [18], [148], [149], [150]. Nonetheless, the prescribed performance control for affine formation maneuver of multi-agent systems is yet to be reported.

Another requirement for formation control is fulfilling control objectives with minimum energy consumption, which can be obtained by building a value function comprising formation tracking errors and control laws [151], [152]. The optimal control laws are derived from the solution of the Hamilton–Jacobi–Bellman equation. Nevertheless, due to the inherent nonlinearity of the Hamilton–Jacobi–Bellman equation, it is difficult to obtain its solution in practice. This problem is usually tackled using neural networks or fuzzy logic systems to compute the approximate solution of the equation via reinforcement learning actor-critic framework. The authors in [153], [154]

deployed reinforcement learning-based optimal control to stir nonlinear agents to track the reference formation geometry. However, the schemes in [153], [154] have some imperfections such as requiring complex algorithms to update the reinforcement learning, persistence excitation must be guaranteed, and the dynamics of the multi-agent systems must be available. In order to relax these conditions, a simplified reinforcement learning optimized nonlinear control based on a simple negative gradient approach has been proposed to solve the tracking problem of strict-feedback nonlinear systems [155], [156], [157], [158], [159], multi-agents consensus problems [160], and formation control problems [161], [162]. All in all, there are a few results about optimal formation control problems based on reinforcement learning. How to formulate an optimal control method using reinforcement learning for leader-follower affine formation maneuvering of nonlinear multi-agent systems is significant, which motivates this chapter.

Inspired by the foregoing discussion, this chapter is focused on the affine formation maneuver control of nonlinear multi-agent systems with prescribed performance subjected to actuator faults and deception attacks. By utilizing the prescribed performance function, an optimized backstepping control coupled with simplified reinforcement learning on identifier-actor-critic architecture is proposed to realize the resilient leader-follower collective affine formation maneuvering in the event of deception attacks and ensure the tracking errors are confined within a predefined region. To the best of our knowledge, this problem has not been addressed. The major contributions of this chapter are summarized as follows:

1. The published results about affine formation maneuver control of multiagent

systems [163], [21], [39], [51], [164], [44], [49], [50], [59], [45], [66], [165] have not addressed the problems of deception attacks or actuator faults but this article takes into account these issues.

2. Even though the affine formation maneuvering of nonlinear multi-agent systems has been studied in [163] using an adaptive proportional-integral-like controller, our work proposes performance-constrained optimized backstepping control with a simple reinforcement learning technique that not only guarantees leader-follower affine formation maneuvering but also resilient to deception attacks and actuator faults.
3. From comparison 3), it is clear that the optimized backstepping control in this article aims to solve a more challenging and rare problem, unlike in [161], [162], [155], [156], [157], [158], [159], [160].
4. Compared with [18], [148], [149], [150], a new prescribed performance function is presented, which can reset to confine the new transient states of the tracking errors within a prespecified set in finite time whenever the reference formation maneuver changes.

6.2 Problem Formulation

Assumption 6.1 [49] *Let $p' = [q_f'^T \ q_l'^T]^T$ be the vector of the target formation maneuvers of the agents, with $q_l' = [q_1'^T \ q_2'^T \ \dots \ q_{N_l}^T]^T$ and $q_f' = [q_{N_l+1}^T \ q_{N_l+2}^T \ \dots \ q_{N_l+N_f}^T]^T$. In this article, it is assumed that the leaders have obtained the desired formation maneuvers*

i.e $q_l = q'_l$ and $u_l = u'_l$. Therefore, the control design for the leaders will be ignored.

The control aim is now to realize $q_f \rightarrow q'_f$ as $t \rightarrow T_s$, with T_s being a finite-time settling time.

6.2.1 Dynamic models of the agents

The leader-follower MAS under consideration consists of N_l leaders and N_f followers.

The leaders are governed by the following dynamic equations:

$$\begin{cases} \dot{q}_i = v_i \\ \dot{v}_i = u_i, \quad i = 1, 2, \dots, N_l \end{cases} \quad (6.1)$$

where $q_i \in \mathbb{R}^2$, $v_i \in \mathbb{R}^2$ and $u_i \in \mathbb{R}^2$, $i = 1, 2, \dots, N_l$ represent the positions, velocities, and control inputs of the leaders, respectively.

The followers are described by second-order nonlinear dynamic equations as follows:

$$\begin{cases} \dot{q}_i = v_i \\ \dot{v}_i = g_i(q_i, v_i)\ddot{u}_i + h_i(q_i, v_i) + \vartheta_i(t) \\ \bar{q}_i = q_i + \chi_i(t, q_i) \\ \bar{v}_i = v_i + \chi_i(t, v_i), \quad i = N_l + 1, \dots, N_l + N_f \end{cases} \quad (6.2)$$

where $q_i \in \mathbb{R}^2$ and $v_i \in \mathbb{R}^2$ represent the positions and velocities of the agents, respectively, $\bar{q}_i \in \mathbb{R}^2$ and $\bar{v}_i \in \mathbb{R}^2$ are the positions and velocities of the agents under deception attacks, $\chi_i(t, q_i)$ and $\chi_i(t, v_i)$ are state-dependent deception attacks sat-

isfying $\chi_i(t, q_i) = w_{q_i} q_i$ and $\chi_i(t, v_i) = w_{v_i} v_i$, with w_{q_i} and w_{v_i} are time-varying and unknown, $u_i \in \mathbb{R}^2$, $\check{u}_i \in \mathbb{R}^2$ represent the faulty actuator, $h_i(q_i, v_i) \in \mathbb{R}^2$ is an unknown continuous nonlinear function, $g_i \in \mathbb{R}^{2 \times 2}$ is a diagonal matrix of unknown input gains, $\vartheta_i(t) \in \mathbb{R}^2$ are the external disturbances.

Remark 6.1 *As a result of the deception attack on the sensors, the original state signals q_i and v_i are not available for control design. Therefore, the corrupted signals from the sensors $\bar{q}_i = (1 + w_{q_i})q_i$ and $\bar{v}_i = (1 + w_{v_i})v_i$ will be utilized for the feedback control design.*

Assumption 6.2 [139] *For the deception attack coefficients $\Lambda_{q_i} = (1 + w_{q_i})$ and $\Lambda_{v_i} = (1 + w_{v_i})$, $1 + w_{q_i} \neq 0$ and $1 + w_{v_i} \neq 0$ throughout the duration of the attack. This assumption means that the system will remain controllable during the attack.*

Let $\Lambda_{q_i} = (1 + w_{q_i})$ and $\Lambda_{v_i} = (1 + w_{v_i})$. Then,

$$\begin{cases} \bar{q}_i = \Lambda_{q_i} q_i \\ \bar{v}_i = \Lambda_{v_i} v_i \end{cases}, \quad i = N_l + 1, \dots, N_l + N_f \quad (6.3)$$

The model of the actuator fault $\check{u}_i$ is described as [145]:

$$\check{u}_i = m_i u_i + \delta_i \quad (6.4)$$

where $u_i \in \mathbb{R}^2$ is the control signal of each agent, and $\delta_i = [\delta_{i1}, \delta_{i2}]^T \in \mathbb{R}^2$ is the vector of bias fault, $m_i = \text{diag}\{m_{i1}, m_{i2}\} \in \mathbb{R}^{2 \times 2}$ is a diagonal matrix of control effectiveness factors, and $0 < m_{i1}, m_{i2} \leq 1$ [145].

Considering (6.3) and (6.4), (6.2) can be expressed as follows:

$$\left\{ \begin{array}{l} \dot{\bar{q}}_i = \Lambda_{q_i} \Lambda_{v_i}^{-1} \bar{v}_i + \dot{\Lambda}_{q_i} \Lambda_{q_i}^{-1} \bar{q}_i \\ \dot{\bar{v}}_i = \Lambda_{v_i} h_i + \Lambda_{v_i} m_i g_i u_i + \Lambda_{v_i} g_i \delta_i + \Lambda_{v_i} \vartheta_i \\ \quad + \dot{\Lambda}_{v_i} \Lambda_{v_i}^{-1} \bar{v}_i, \quad i = N_l + 1, \dots, N_l + N_f \end{array} \right. \quad (6.5)$$

Let's define

$$q_f = [\bar{q}_{N_l+1}^T \ \bar{q}_{N_l+2}^T \ \dots \ \bar{q}_{N_l+N_f}^T]^T$$

$$v_f = [\bar{v}_{N_l+1}^T \ \bar{v}_{N_l+2}^T \ \dots \ \bar{v}_{N_l+N_f}^T]^T$$

$$h = [h_{N_l+1}^T \ h_{N_l+2}^T \ \dots \ h_{N_l+N_f}^T]^T$$

$$\vartheta = [\vartheta_{N_l+1}^T \ \vartheta_{N_l+2}^T \ \dots \ \vartheta_{N_l+N_f}^T]^T$$

$$\delta = [\delta_{N_l+1}^T \ \delta_{N_l+2}^T \ \dots \ \delta_{N_l+N_f}^T]^T$$

$$\Lambda_p = \text{diag}\{\Lambda_{q_{N_l+1}} \ \Lambda_{q_{N_l+2}} \ \dots \ \Lambda_{q_{N_l+N_f}}\}$$

$$\Lambda_v = \text{diag}\{\Lambda_{v_{N_l+1}} \ \Lambda_{v_{N_l+2}} \ \dots \ \Lambda_{v_{N_l+N_f}}\}$$

$$m = \text{diag}\{m_{N_l+1} \ m_{N_l+2} \ \dots \ m_{N_l+N_f}\}$$

$$g = \text{diag}\{g_{N_l+1} \ g_{N_l+2} \ \dots \ g_{N_l+N_f}\}$$

The global form of (6.5) is thus:

$$\left\{ \begin{array}{l} \dot{q}_f = \Lambda_q \Lambda_v^{-1} v_f + \dot{\Lambda}_q \Lambda_q^{-1} q_f \\ \dot{v}_f = \Lambda_v m g u + \Lambda_v \delta + \Lambda_v h + \Lambda_v \vartheta + \dot{\Lambda}_v \Lambda_v^{-1} v_f \end{array} \right. \quad (6.6)$$

Define the following error variables for the followers:

$$\varepsilon_i = \sum_{j=1}^N \omega_{ij}(q_i - q_j), \quad i = N_l + 1, \dots, N_l + N_f \quad (6.7)$$

where ω_{ij} is the stress weight as defined in (2.4.1).

and the compact form of (6.7) is given by [21]:

$$\begin{aligned} \varepsilon_f &= \bar{\Omega}_{ff} q_f + \bar{\Omega}_{fl} q_l = \bar{\Omega}_{ff} \left[q_f - (-\Omega_{ff}^{-1} \Omega_{fl} \otimes I_d) q_l \right] \\ &= \bar{\Omega}_{ff} \left[q_f - q'_f \right] \end{aligned} \quad (6.8)$$

where $\varepsilon_f = [\varepsilon_1^T \ \varepsilon_2^T \ \dots \ \varepsilon_{N_f}^T]^T$, and $q'_f = (-\Omega_{ff}^{-1} \Omega_{fl} \otimes I_d) q_l$ as in (2.10).

6.2.2 Neural network approximations

In this paper, the smooth and continuous nonlinear functions derived from the design of the proposed control strategy are approximated with radial basis function neural networks (RBFNNs) as follows:

$$\Xi_i = W_i^T \Phi_i(X_i) + \epsilon_i(X_i) \quad (6.9)$$

where $W_i = [w_{i1} \ w_{i2} \ \dots \ w_{im}]^T \in \mathbb{R}^m$ is the weight vector, m is the number of nodes, $\epsilon_i(X_i)$ is the RBFNN reconstruction error, error satisfying $\|\epsilon_i(X_i)\| \leq \bar{\epsilon}_i$, $\Phi_i(X_i) = [\Phi_{i1}(X_i) \ \Phi_{i2}(X_i) \ \dots \ \Phi_{im}(X_i)]^T$ is the vector of basis function with:

$$\Phi_{ik}(X_i) = \exp \left(-\frac{\|X_i - \eta_{1k}\|^2}{\eta_{2k}^2} \right), \quad k = 1, 2, \dots, m \quad (6.10)$$

where η_{1k} is the receptive center of the Gaussian function, η_{2k} is the Gaussian function width.

6.2.3 Finite Time Prescribed Performance Function

The specified performance is attained by confining each leader-follower tracking error ε_i , $i = N_l + 1, \dots, N_l + N_f$ within the following preassigned boundary.

$$-\underline{\delta}_i \zeta_i(t) < \varepsilon_i < \bar{\delta}_i \zeta_i(t) \quad (6.11)$$

where $\underline{\delta}_i > 0$ and $\bar{\delta}_i$ are design constants, $\zeta_i(t) : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is the prescribed performance function characterized as a positive, smooth, and decreasing function.

Definition 6.1 [166] *A smooth function $\zeta_i(t) : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is said to be a finite time prescribed performance function if it is characterized by a) $\zeta_i(t) > 0$; b) $\dot{\zeta}_i(t) < 0$; c) $\lim_{t \rightarrow t_s} \zeta_i(t) = \zeta_{is} > 0$; and d) For any $t \geq t_s$, $\zeta_i(t) = \zeta_{is}$.*

The finite time prescribed performance function proposed in this study is defined by (6.12):

$$\zeta_i(t) = \begin{cases} \zeta_{i0}, & t = t_n \\ (\zeta_{i0} - \zeta_{is}) \exp\left(\frac{-\kappa_i(t-t_n)}{T_s}\right) + \zeta_{is}, & 0 < t - t_n < T_s \\ \zeta_{is}, & t - t_n \geq T_s \end{cases}$$

$$i = N_l + 1, \dots, N_l + N_f \quad (6.12)$$

where $\zeta_{i0} > 0$, $\zeta_{is} > 0$, and $\kappa_i > 0$ are appropriately selected constants, T_s is the preassigned finite settling time, and t_n is the time at which the agents undergo new transient states due to the change in target formation maneuver.

The inequality (6.11) can be transformed into an equality form using the following error transformation:

$$e_i = \Gamma \left(\frac{\varepsilon_i}{\zeta_i} \right) \quad (6.13)$$

where e_i is the transformed error, $\Gamma(\cdot)$ is a smooth function and $\Gamma : (-\underline{\delta}_i, \bar{\delta}_i) \rightarrow (-\infty, \infty)$. The transformation function $\Gamma(\cdot)$ is expressed as:

$$\Gamma \left(\frac{\varepsilon_i}{\zeta_i} \right) = \frac{1}{2} \ln \left(\frac{\underline{\delta}_i + \frac{\varepsilon_i}{\zeta_i}}{\bar{\delta}_i - \frac{\varepsilon_i}{\zeta_i}} \right) \quad (6.14)$$

The time derivative of the transformed error yields

$$\dot{e}_i = \hbar_i \left(\dot{\varepsilon}_i - \frac{\dot{\zeta}_i}{\zeta_i} \varepsilon_i \right) \quad i = N_l + 1, \dots, N_l + N_f \quad (6.15)$$

where $\hbar_i = \frac{1}{\zeta_i} \frac{\partial \Gamma_i}{\partial (\varepsilon_i / \zeta_i)}$.

To better demonstrate the important features of the proposed performance function, a simple example is presented in Fig. 6.1. This figure shows a state trajectory q_i tracking a target trajectory q'_i with sudden change in amplitudes at $t_n = 4$ s and $t_n = 8$ s. The trajectory q_i experienced new transient states at the time $t = t_n = 4$ s and $t = t_n = 8$ s. It is required that q_i settles at T_s seconds after every new transient state. The finite time performance function (6.12) is able reset and constrained the new transient states of the tracking error ε_i whenever the amplitude of the target

trajectory changes.

- Initially, $t = t_n = 0$ and $\zeta_i(0) = \zeta_{i0} = 6$. It is expected that as t grows, $\zeta_i(t)$ converges to $\zeta_{is} = 0.3$ at $t = T_s = 0.5$ s. Then, $\zeta_i(t) = \zeta_{is}$ for $t \geq T_s$.
- When there is a sudden decrease in the target trajectory at $t = t_n = 4$ s, $t - t_n = 0$ and the finite time performance function reset to $\zeta_i(0) = \zeta_{i0}$ to confine the new transient state within the predefined range. As t grows (i.e $t - 4 > 0$), the finite time performance function decays and settles at $t - 4 = 0.5$ s. For $t - 4 \geq 0.5$ s, $\zeta_i(t) = \zeta_{is}$.
- On the other hand, when the target trajectory suddenly increases at $t = t_n = 8$ s, $t - t_n = 0$. At this moment, the finite time performance function reset such that $\zeta_i(0) = \zeta_{i0}$ and begins decaying for $t - 8 > 0$ until $t - 8 = 0.5$ s where it settles at $\zeta_i(t) = \zeta_{is}$. Subsequently, for $t - 8 \geq 0.5$ s, $\zeta_i(t) = \zeta_{is}$.

Remark 6.2 Various performance functions that differ from (6.12) have been developed in the literature including (6.16) [150], (6.17) [167], (6.18) [168], (6.19) [169], (6.20) [170], (6.21) [171], and (6.22) [172], and (6.23) [166]

$$\zeta_i(t) = (\zeta_{i0} - \zeta_{i\infty}) e^{-\kappa_i t} + \zeta_{i\infty} \quad (6.16)$$

$$\zeta_i(t) = \zeta_{i0} \exp\left(\frac{-t}{\kappa_i}\right) + \zeta_{i\infty} \quad (6.17)$$

$$\zeta_i(t) = \coth(\kappa_{1i}t + \kappa_{i2}) - 1 + \zeta_{i\infty} \quad (6.18)$$

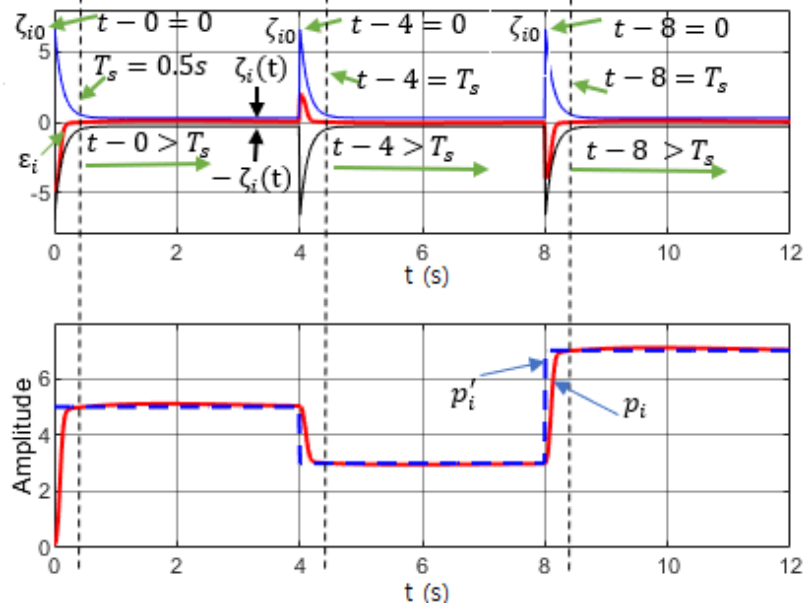


Figure 6.1: A simple example showing the performance of the proposed performance function.

$$\zeta_i(t) = \begin{cases} (\zeta_{i0} - \zeta_{is}) \exp\left(\frac{-\kappa_i T_s t}{t - T_s}\right) + \zeta_{is}, & t < T_s \\ \zeta_{is}, & t \geq T_s \end{cases} \quad (6.19)$$

$$\zeta_i(t) = \begin{cases} \zeta_{i0} \exp\left(\kappa_i \left(1 - \frac{T_s}{T_s - t}\right)\right) \varpi_i + \zeta_{is}, & t < T_s \\ \varpi_i = \frac{1}{2} \cos\left(\frac{\pi t}{T_s}\right) + \frac{1}{2} & \\ \zeta_{is}, & t \geq T_s \end{cases} \quad (6.20)$$

$$\zeta_i(t) = \begin{cases} \left(1 - \frac{t}{T_s}\right) \exp^{1 - \exp\left(\frac{T_s}{T_s - t}\right)} + \zeta_{is}, & t < T_s \\ \zeta_{is}, & t \geq T_s \end{cases} \quad (6.21)$$

$$\zeta_i(t) = \begin{cases} \left(\zeta_{i0} - \frac{t}{\kappa_i T_s} \right) \\ \quad \times \exp \left(1 - \tan \left(\frac{t}{T_s - t} \right) \right) + \zeta_{is}, & t < T_s \\ \zeta_{is}, & t \geq T_s \end{cases} \quad (6.22)$$

$$\zeta_i(t) = \begin{cases} (\zeta_{i0} - \zeta_{is}) \left(1 - \frac{t}{T_s} \right)^{\kappa_i} + \zeta_{is}, & t < T_s \\ \zeta_{is}, & t \geq T_s \end{cases} \quad (6.23)$$

Remark 6.3 *It is obvious that (6.16), (6.17), and (6.18) cannot converge to the pre-defined steady-state within a specified finite time while (6.19), (6.20), (6.21), (6.22), and (6.23) can converge within a specified settling time. However, (6.19), (6.20), (6.21), and (6.22) have the problem of singularity at $t = T_s$. Moreover, the term $\left(1 - \frac{t}{T_s}\right)$ in (6.23) is a first-order approximation of the Taylor-series expansion of $\exp\left(-\frac{t}{T_s}\right) = 1 - \frac{t}{T_s} + 1/2 \frac{t^2}{T_s^2} - 1/6 \frac{t^3}{T_s^3} + 1/24 \frac{t^4}{T_s^4} + \dots$, and therefore (6.23) is less accurate. The essential drawback of (6.16)-(6.23) is that when there is an abrupt change in the reference signal they cannot reset to accommodate the new transient state of the tracking errors. Moreover, they cannot be employed for application in affine formation maneuvering of multi-agent systems where the dynamics of the mobile agents undergo new transient states before converging to the new formation maneuver. This motivates us to formulate a modified prescribed performance function that can reset at any time to accommodate the new transient states of the tracking errors, and then converge to the steady-state in a prescribed finite time.*

Fig. 6.2 depicts the limitation of (6.16)-(6.23) when there is a sudden change

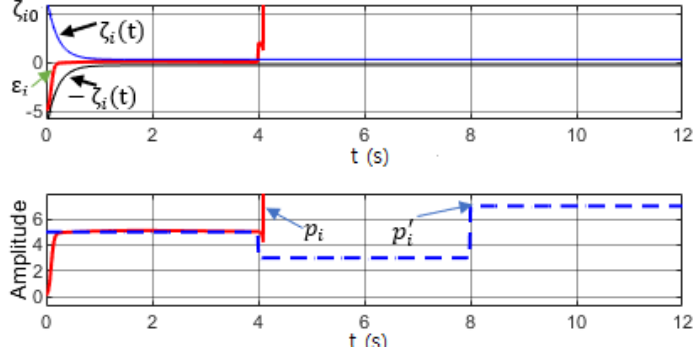


Figure 6.2: A simple example showing the performance of existing performance functions.

in the target trajectory. It can be seen that the trajectory q_i becomes unstable at $t = t_n = 4$ s because the existing performance functions cannot reset to take care of the new transient state of the tracking error. Therefore, the existing performance functions cannot be used in applications where the agents undergo multiple transient states.

6.3 Control design

This section proposes a reinforcement learning-based optimized secure backstepping control for nonlinear leader-follower multi-agent systems with deception attacks and actuator faults to be resilient and realize various affine formation maneuvers. The main aim of the proposed controller is to guarantee that the closed-loop system is bounded despite the actuator and the deceptive state signals injected by cyber-attackers.

Lets define:

$$\begin{aligned}
e_f &= [e_1^T \ e_2^T \ \dots \ e_{N_f}^T]^T \\
\zeta_f &= \text{diag}\{\dot{\zeta}_1/\zeta_1 \ \dot{\zeta}_2/\zeta_2 \ \dots \ \dot{\zeta}_{N_f}/\zeta_{N_f}\} \\
\hbar_f &= \text{diag}\{\hbar_1 \ \hbar_2 \ \dots \ \hbar_{N_f}\}
\end{aligned}$$

The global form of (6.15) is given by:

$$\begin{aligned}
\dot{e}_f &= \hbar_f (\dot{\varepsilon}_f - \zeta_f \varepsilon_f) \\
&= \hbar_f \bar{\Omega}_{ff} \left(\Lambda_q \Lambda_v^{-1} v_f + \dot{\Lambda}_q \Lambda_q^{-1} q_f - \dot{q}'_f \right) - \zeta_f \hbar_f \varepsilon_f
\end{aligned} \tag{6.24}$$

According to the backstepping approach, v_f will be treated as the intermediate control input. Let α_f be the virtual control. Then, define the error $z_f = v_f - \alpha_f$, and its time-derivative is calculated as:

$$\begin{aligned}
\dot{z}_f &= \dot{v}_f - \dot{\alpha}_f \\
&= \Lambda_v m g u + \Lambda_v g \delta + \Lambda_v h + \Lambda_v \vartheta + \dot{\Lambda}_v \Lambda_v^{-1} v_f - \dot{\alpha}_f
\end{aligned} \tag{6.25}$$

For the purpose of designing the proposed control scheme, the error dynamics (6.24) and (6.25) are transformed as follows [173], [174]:

$$\begin{aligned}
\dot{e}_f &= \Upsilon(q_f, v_f, \Lambda_q, \Lambda_v) + v_f \\
&= \Upsilon(q_f, v_f, \Lambda_q, \Lambda_v) + z_f + \alpha_f
\end{aligned} \tag{6.26}$$

$$\dot{z}_f = F(q_f, v_f, \Lambda_q, \Lambda_v, u) + u - \dot{\alpha}_f \quad (6.27)$$

where $\Upsilon(q_f, v_f, \Lambda_q, \Lambda_v) = \hbar_f \bar{\Omega}_{ff} \left(\Lambda_q \Lambda_v^{-1} v_f + \dot{\Lambda}_q \Lambda_q^{-1} q_f - \dot{q}'_f \right) - \zeta_f \hbar_f \varepsilon_f - v_f$ and $F(q_f, v_f, \Lambda_q, \Lambda_v, u) = \Lambda_v m g u + \Lambda_v g \delta + \Lambda_v h + \Lambda_v \vartheta + \dot{\Lambda}_v \Lambda_v^{-1} v_f - u$.

Remark 6.4 *By transforming (6.24) and (6.25) into (6.26) and (6.27), respectively, the problems of the unknown control coefficients are avoided. So, there is no need to use the Nussbaum gain function [135]. Moreover, the common assumption that the unknown input gain matrix g or the unknown control effectiveness matrix m is bounded by known upper bounds has been relaxed. The unknown functions $\Upsilon(q_f, v_f, \Lambda_q, \Lambda_v)$ and $F(q_f, v_f, \Lambda_q, \Lambda_v, u)$ will be approximated with the aid of neural networks.*

A performance index function associated with e_f and α_f is defined as follows:

$$I_p(e_f) = \int_t^\infty r_f(e_f(\tau), \alpha_f(\tau)) d\tau \quad (6.28)$$

where $r_f(e_f, \alpha_f) = e_f^T e_f + \alpha_f^T \alpha_f$ is the cost function. The optimal performance index $I^*(e_f)$ associated with the optimal virtual controller α_f^* is given by:

$$\begin{aligned} I_p^*(e_f) &= \min_{\alpha_f \in \Pi} \left(\int_t^\infty r_f(e_f(\tau), \alpha_f(\tau)) d\tau \right) \\ &= \int_t^\infty (e_f^T e_f + \alpha_f^{*T} \alpha_f^*) d\tau \end{aligned} \quad (6.29)$$

where Π is the set of admissible control inputs. From (6.29), the following HJB

equation can be derived:

$$\begin{aligned}
H_p\left(e_f, \alpha_f^*, \frac{\partial I_p^*}{\partial e_f}\right) &= r_f(e_f, \alpha_f^*) + \frac{\partial I_p^*}{\partial e_f} \dot{e}_f = e_f^T e_f \\
&+ \alpha_f^{*T} \alpha_f^* + \frac{\partial I_p^*}{\partial e_f} (z_f + \alpha_f + \Upsilon - q_f') = 0
\end{aligned} \tag{6.30}$$

where $\frac{\partial I_p^*}{\partial e_f}$ is the gradient of I_p^* along e_f . Suppose the solution of (6.30) exists and is unique, the optimal control policy α_f^* can be achieved by computing $\partial H_p\left(e_f, \alpha_f^*, \frac{\partial I_p^*}{\partial e_f}\right) / \partial \alpha_f^* = 0$

$$\alpha_f^* = -\frac{1}{2} \frac{\partial I_q^*(e_f)}{\partial e_f} \tag{6.31}$$

Considering (6.30) and (6.31), it is clear that $\frac{\partial I_p^*}{\partial e_f}$ is required to obtain the solution of (6.30). Nevertheless, due to the unknown deception attack signals and strong nonlinearities, it is impossible to solve (6.30). To achieve the control objective, RL actor-critic framework is employed to obtain the approximate solution online.

By using some mathematical manipulations, $\frac{\partial I_p^*}{\partial e_f}$ can be expressed as follows:

$$\frac{\partial I_q^*(e_f)}{\partial e_f} = 2K_1 e_f + 2\Upsilon + I_p^0 \tag{6.32}$$

where $I_p^0 = -2K_1 e_f - 2\Upsilon + \frac{\partial I_q^*(e_f)}{\partial e_f}$.

Inserting (6.32) into (6.31) gives:

$$\alpha_f^* = -K_p e_f - \Upsilon - \frac{1}{2} I_q^0 \tag{6.33}$$

Since the continuous functions Υ and I_q^0 are unknown, RBFNNs are employed to

approximate them in the following form:

$$\Upsilon = W_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) + \epsilon_{\Upsilon}(X_{\Upsilon}) \quad (6.34)$$

$$I_q^0 = W_{I_q}^{*T} \Phi_{I_q}(X_{I_q}) + \epsilon_{I_q}(X_{I_q}) \quad (6.35)$$

where W_{Υ}^* and $W_{I_q}^*$ are the ideal weights, $\Phi_{\Upsilon}(X_{\Upsilon})$ and $\Phi_{I_q}(X_{I_q})$ are the vectors basis functions, and $\epsilon_{\Upsilon}(X_{\Upsilon})$ and $\epsilon_{I_q}(X_{I_q})$ are the RBFNN approximation errors.

Substituting (6.34) and (6.35) into (6.32) and (6.33), respectively yields:

$$\begin{aligned} \frac{\partial I_q^*(e_f)}{\partial e_f} &= 2K_1 e_f + 2W_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ &\quad + W_{I_q}^{*T} \Phi_{I_q}(X_{I_q}) + \epsilon_p \end{aligned} \quad (6.36)$$

$$\begin{aligned} \alpha_f^* &= -K_p e_f - W_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ &\quad - \frac{1}{2} W_{I_q}^{*T} \Phi_{I_q}(X_{I_q}) + \epsilon_p \end{aligned} \quad (6.37)$$

where $\epsilon_p = \epsilon_{\Upsilon}(X_{\Upsilon}) + \epsilon_{I_q}(X_{I_q})$.

The optimal control input (6.37) cannot be used because of the unknown weights W_{Υ}^* and $W_{I_q}^*$. The identifier RBFNN is used to estimate the unknown deceptive signals as:

$$\hat{\Upsilon} = \hat{W}_{\Upsilon}^T \Phi_{\Upsilon}(X_{\Upsilon}) \quad (6.38)$$

where $\hat{W}_{\Upsilon}$ is the weight of the identifier, and $\hat{\Upsilon}$ is the output of the identifier. The RBFNN weight of the identifier $\hat{W}_{\Upsilon}^T$ is updated online by:

$$\dot{\hat{W}}_{\Upsilon} = \Gamma_q \left(\Phi_{\Upsilon}(X_{\Upsilon}) e_f - \sigma_q \hat{W}_{\Upsilon} \right) \quad (6.39)$$

where Γ_q is a positive-definite matrix and $\sigma_q > 0$ is a small constant.

To obtain the usable optimized controller, the actor, critic, and identifier framework is developed:

$$\frac{\partial \hat{I}_q(e_f)}{\partial e_f} = 2K_1 e_f + 2\hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) + \hat{W}_{c_q}^{*T} \Phi_{I_q}(X_{I_q}) \quad (6.40)$$

$$\hat{\alpha}_f^* = -K_p e_f - \hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) - \frac{1}{2} \hat{W}_{a_q}^{*T} \Phi_{I_q}(X_{I_q}) \quad (6.41)$$

where $\frac{\partial \hat{I}_q(e_f)}{\partial e_f}$ is the estimate of $\frac{\partial I_q(e_f)}{\partial e_f}$, $\hat{W}_{c_p}^*$ and $\hat{W}_{a_q}^*$ are the estimates of the critic and actor weights, respectively.

Adding (6.40) and (6.41) into (6.30), the approximated HJB equation is derived as:

$$\begin{aligned} H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right) &= e_f^T e_f \\ &+ \begin{pmatrix} -K_p e_f - \hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ -\frac{1}{2} \hat{W}_{a_q}^{*T} \Phi_{I_q}(X_{I_q}) \end{pmatrix}^T \\ &\times \begin{pmatrix} -K_p e_f - \hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ -\frac{1}{2} \hat{W}_{a_q}^{*T} \Phi_{I_q}(X_{I_q}) \end{pmatrix} \\ &+ \begin{pmatrix} 2K_1 e_f + 2\hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ +\hat{W}_{c_q}^{*T} \Phi_{I_q}(X_{I_q}) \end{pmatrix} \\ &\times \begin{pmatrix} z_f + \Upsilon \\ -K_p e_f - \hat{W}_{\Upsilon}^{*T} \Phi_{\Upsilon}(X_{\Upsilon}) \\ -\frac{1}{2} \hat{W}_{a_q}^{*T} \Phi_{I_q}(X_{I_q}) \end{pmatrix} \end{aligned} \quad (6.42)$$

The Bellman's residual error $\tilde{H}_p$ is expressed as:

$$\tilde{H}_p = H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right) - H_p\left(e_f, \alpha_f^*, \frac{\partial I_p^*}{\partial e_f}\right) \quad (6.43)$$

Considering (6.30), (6.43) becomes:

$$\tilde{H}_p = H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right) \quad (6.44)$$

With regards to (6.42), it is desired that $\hat{\alpha}_f^*$ realizes $\tilde{H}_p = H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right) \rightarrow 0$.

If $H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right) = 0$ ensued, the following also hold:

$$\frac{\partial_f H_p\left(e_f, \hat{\alpha}_f^*, \frac{\partial \hat{I}_p}{\partial e_f}\right)}{\partial \hat{W}_{a_q}} = \frac{1}{2} \Phi_{I_q}^T \Phi_{I_q} (\hat{W}_{a_q} - \hat{W}_{c_q}) = 0 \quad (6.45)$$

To derive the weight updating laws for $\hat{W}_{a_q}$ and $\hat{W}_{c_q}$ to guarantee (6.45), a positive definite function S_p is defined as:

$$S_p = (\hat{W}_{a_q} - \hat{W}_{c_q})^T (\hat{W}_{a_q} - \hat{W}_{c_q}) \quad (6.46)$$

Clearly, $S_p = 0$ is equivalent to (6.45). Noting that $\partial S_q / \partial \hat{W}_{a_q} = -\partial S_q / \partial \hat{W}_{c_p} = 2(\hat{W}_{a_q} - \hat{W}_{c_p})$, the time-derivative of S_p gives

$$\begin{aligned} \frac{dS_q}{dt} &= \frac{\partial S_q}{\partial \hat{W}_{c_p}^T} \dot{\hat{W}}_{c_p} + \frac{\partial S_q}{\partial \hat{W}_{a_p}^T} \dot{\hat{W}}_{a_p} = -k_c \frac{\partial S_q}{\partial \hat{W}_{c_p}^T} \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \\ &\quad - \frac{\partial S_q}{\partial \hat{W}_{a_p}^T} \Phi_{I_q}^T \Phi_{I_q} \left(k_a (\hat{W}_{a_p} - \hat{W}_{c_p}) + k_c \hat{W}_{c_p} \right) \\ &= -\frac{k_a}{2} \frac{\partial S_q}{\partial \hat{W}_{a_p}^T} \Phi_{I_q}^T \Phi_{I_q} \frac{\partial S_q}{\partial \hat{W}_{a_i}} \leq 0 \end{aligned} \quad (6.47)$$

The RBFNN weight of the critic is updated as follows:

$$\dot{\hat{W}}_{c_p} = -\gamma_{c_p} \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \quad (6.48)$$

where $\gamma_{c_p} > 0$ is a constant.

The RBFNN weight of the actor is updated as follows:

$$\dot{\hat{W}}_{a_p} = -\Phi_{I_q}^T \Phi_{I_q} \left(\gamma_{a_p} \left(\hat{W}_{a_p} - \hat{W}_{c_p} \right) + \gamma_{c_p} \hat{W}_{c_p} \right) \quad (6.49)$$

where $\gamma_{a_p} > 0$ is a constant.

Using the approximate optimal controller (6.41), the error dynamics (6.26) can be rewritten as:

$$\dot{e}_f = z_f + \hat{\alpha}_f^* + \Upsilon \quad (6.50)$$

A Lyapunov candidate function is selected for the e_f -subsystem as:

$$\begin{aligned} L_1 = & \frac{1}{2} e_f^T e_f + \frac{1}{2} \tilde{W}_\Upsilon^T \Gamma_q^{-1} \tilde{W}_\Upsilon \\ & + \frac{1}{2} \tilde{W}_{c_p}^T \tilde{W}_{c_p} + \frac{1}{2} \tilde{W}_{a_p}^T \tilde{W}_{a_p} \end{aligned} \quad (6.51)$$

where $\tilde{W}_\Upsilon = \hat{W}_\Upsilon - W_\Upsilon^*$, $\tilde{W}_{a_p} = \hat{W}_{a_p} - W_{I_p}^*$, and $\tilde{W}_{c_p} = \hat{W}_{c_p} - W_{I_p}^*$ are the RBFNN weight errors for the identifier, actor, and critic networks, respectively.

Differentiating L_1 with respect to time and considering (6.39), (6.48), and (6.49), one gets:

$$\begin{aligned}
\dot{L}_1 = & e_f \left(z_f + \hat{\alpha}_f^* + \Upsilon \right) + \tilde{W}_\Upsilon^T \left(\Phi_\Upsilon^T e_f - \sigma_p \hat{W}_\Upsilon \right) \\
& - \gamma_{c_p} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \\
& - \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \left(\gamma_{a_p} \left(\hat{W}_{a_p} - \hat{W}_{c_p} \right) + \gamma_{c_p} \hat{W}_{c_p} \right)
\end{aligned} \tag{6.52}$$

Inserting (6.34) and (6.41) into (6.52), one gets:

$$\begin{aligned}
\dot{L}_1 = & e_f \left(-K_p e_f - \frac{1}{2} \hat{W}_{a_q}^T \Phi_{I_q} (X_{I_q}) \right. \\
& \left. - \tilde{W}_\Upsilon^T \Phi_\Upsilon (X_\Upsilon) + \epsilon_\Upsilon + z_f \right) \\
& + \tilde{W}_\Upsilon^T \left(\Phi_\Upsilon^T e_f - \sigma_p \hat{W}_\Upsilon \right) - \gamma_{c_p} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \\
& - \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \left(\gamma_{a_p} \left(\hat{W}_{a_p} - \hat{W}_{c_p} \right) + \gamma_{c_p} \hat{W}_{c_p} \right)
\end{aligned} \tag{6.53}$$

Equation (6.53) can be re-expressed as follows:

$$\begin{aligned}
\dot{L}_1 = & -e_f K_p e_f + e_f \epsilon_\Upsilon + e_f z_f - \frac{1}{2} e_f^T \hat{W}_{a_q}^T \Phi_{I_q} \\
& - \sigma_p \tilde{W}_\Upsilon^T \hat{W}_\Upsilon - \gamma_{c_p} \tilde{W}_{c_p}^T \Phi_{I_q}^T \hat{h}_f \Omega_{ff} \Phi_{I_q} \hat{W}_{c_p} \\
& - \tilde{W}_{a_p}^T \Phi_{I_q}^T \hat{h}_f \Omega_{ff} \Phi_{I_q} \\
& \times \left(\gamma_{a_p} \left(\hat{W}_{a_p} - \hat{W}_{c_p} \right) + \gamma_{c_p} \hat{W}_{c_p} \right)
\end{aligned} \tag{6.54}$$

By utilizing Young's inequality, we have:

$$\begin{aligned}
e_f^T \epsilon_\Upsilon & \leq \frac{1}{2} e_f^T e_f + \frac{1}{2} \epsilon_\Upsilon^T \epsilon_\Upsilon \\
e_f^T z_f & \leq \frac{1}{2} e_f^T e_f + \frac{1}{2} z_f^T z_f \\
-\frac{1}{2} e_f^T \hat{W}_{a_q}^T \Phi_{I_q} & \leq \frac{1}{4} e_f^T e_f + \frac{1}{4} \hat{W}_{a_q}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{a_q}
\end{aligned} \tag{6.55}$$

Substituting the Young's inequality (6.55) into (6.54) gives:

$$\begin{aligned}
\dot{L}_1 = & -e_f(K_p - 1.25)e_f - \sigma_p \tilde{W}_\Upsilon^T \hat{W}_\Upsilon \\
& - \gamma_{c_p} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} - \gamma_{a_p} \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{a_p} \\
& + (\gamma_{a_p} - \gamma_{a_c}) \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \\
& + \frac{1}{4} \hat{W}_{a_q}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{a_q} + \frac{1}{2} \epsilon_\Upsilon^T \epsilon_\Upsilon + \frac{1}{2} z_f^T z_f
\end{aligned} \tag{6.56}$$

Noting that $\tilde{W}_\Upsilon = \hat{W}_\Upsilon - W_\Upsilon^*$, $\tilde{W}_{a_p} = \hat{W}_{a_p} - W_{I_p}^*$, and $\tilde{W}_{c_p} = \hat{W}_{c_p} - W_{I_p}^*$, the following facts are valid:

$$\tilde{W}_\Upsilon^T \hat{W}_\Upsilon = \frac{1}{2} \tilde{W}_\Upsilon^T \tilde{W}_\Upsilon + \frac{1}{2} \hat{W}_\Upsilon^T \hat{W}_\Upsilon - \frac{1}{2} W_\Upsilon^{*T} W_\Upsilon^* \tag{6.57}$$

$$\begin{aligned}
\tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} = & \frac{1}{2} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{c_p} + \frac{1}{2} \hat{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{c_p} \\
& - \frac{1}{2} W_{I_p}^{*T} \Phi_{I_q}^T \Phi_{I_q} W_{I_p}^*
\end{aligned} \tag{6.58}$$

$$\begin{aligned}
\tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{a_p} = & \frac{1}{2} \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{a_p} + \frac{1}{2} \hat{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \hat{W}_{a_p} \\
& - \frac{1}{2} W_{I_p}^{*T} \Phi_{I_q}^T \Phi_{I_q} W_{I_p}^*
\end{aligned} \tag{6.59}$$

Combining (6.57)-(6.59) with (6.56) yields:

$$\begin{aligned}
\dot{L}_1 = & -e_f(K_p - 1.25)e_f - \frac{\sigma_p}{2}\tilde{W}_{\Upsilon}^T\tilde{W}_{\Upsilon} \\
& - \frac{\gamma_{c_p}}{2}\tilde{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{c_p} - \frac{\gamma_{a_p}}{2}\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{a_p} \\
& + (\gamma_{a_p} - \gamma_{a_c})\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{c_p} - \frac{\gamma_{c_p}}{2}\hat{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{c_p} \\
& - \left(\frac{\gamma_{a_p}}{2} - \frac{1}{4}\right)\hat{W}_{a_q}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{a_q} \\
& + \left(\frac{\gamma_{a_p}}{2} + \frac{\gamma_{c_p}}{2}\right)W_{I_p}^{*T}\Phi_{I_q}^T\Phi_{I_q}W_{I_p}^* \\
& + \frac{\sigma_p}{2}W_{\Upsilon}^{*T}W_{\Upsilon}^* + \frac{1}{2}\epsilon_{\Upsilon}^T\epsilon_{\Upsilon} + \frac{1}{2}z_f^Tz_f
\end{aligned} \tag{6.60}$$

According to Young's inequality, the following relationship is valid

$$\begin{aligned}
\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{c_p} &= \frac{1}{2}\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{a_p} \\
&+ \frac{1}{2}\hat{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{c_p}
\end{aligned} \tag{6.61}$$

then

$$\begin{aligned}
\dot{L}_1 = & -e_f(K_p - 1.25)e_f - \frac{\sigma_p}{2}\tilde{W}_{\Upsilon}^T\tilde{W}_{\Upsilon} \\
& - \frac{\gamma_{c_p}}{2}\tilde{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{c_p} - \frac{\gamma_{c_p}}{2}\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{a_p} \\
& - \left(\gamma_{c_p} - \frac{\gamma_{a_p}}{2}\right)\hat{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{c_p} \\
& - \left(\frac{\gamma_{a_p}}{2} - \frac{1}{4}\right)\hat{W}_{a_q}^T\Phi_{I_q}^T\Phi_{I_q}\hat{W}_{a_q} \\
& + \left(\frac{\gamma_{a_p}}{2} + \frac{\gamma_{c_p}}{2}\right)W_{I_p}^{*T}\Phi_{I_q}^T\Phi_{I_q}W_{I_p}^* \\
& + \frac{\sigma_p}{2}W_{\Upsilon}^{*T}W_{\Upsilon}^* + \frac{1}{2}\epsilon_{\Upsilon}^T\epsilon_{\Upsilon} + \frac{1}{2}z_f^Tz_f
\end{aligned} \tag{6.62}$$

The parameters K_p , γ_{c_p} , and γ_{a_p} are selected such that:

$$K_p > 1.25, \gamma_{a_p} > \frac{1}{2}, \gamma_{a_p} > \gamma_{c_p} > \frac{\gamma_{a_p}}{2} \quad (6.63)$$

Therefore, based on the selected parameters, (6.62) becomes:

$$\begin{aligned} \dot{L}_1 = & -e_f(K_p - 2)e_f - \frac{\sigma_p}{2}\tilde{W}_\Upsilon^T\tilde{W}_\Upsilon - \frac{\gamma_{c_p}}{2}\tilde{W}_{c_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{c_p} \\ & - \frac{\gamma_{c_p}}{2}\tilde{W}_{a_p}^T\Phi_{I_q}^T\Phi_{I_q}\tilde{W}_{a_p}\frac{1}{2}z_f^Tz_f + \mathcal{D}_1 \end{aligned} \quad (6.64)$$

where $\mathcal{D}_1 = \left(\frac{\gamma_{a_p}}{2} + \frac{\gamma_{c_p}}{2}\right)W_{I_p}^{*T}\Phi_{I_q}^T\Phi_{I_q}W_{I_p}^* + \frac{\sigma_p}{2}W_\Upsilon^{*T}W_\Upsilon^* + \frac{1}{2}\epsilon_\Upsilon^T\epsilon_\Upsilon$, which satisfies $\mathcal{D}_1 \leq \bar{\mathcal{D}}_1$ and $\bar{\mathcal{D}}_1 > 0$ is a constant.

After designing the optimized virtual control α_f^* , the next step is designing the actual control u . Equation (6.27) can be rewritten with the approximate virtual controller $\hat{\alpha}_f^*$ as follows:

$$\dot{z}_f = F + u - \hat{\alpha}_f^* \quad (6.65)$$

A performance index function associated with z_f and u is defined as follows:

$$I_v(z_f) = \int_t^\infty r_v(z_f(\tau), u(\tau))d\tau \quad (6.66)$$

where $r_v(z_f, u) = z_f^T z_f + u^T u$ is the cost function. Let u^* be the optimal control and

the corresponding optimal performance index $I(z_f)^*$ is constructed as:

$$\begin{aligned} I_v^*(z_f) &= \min_{u \in \Pi} \left(\int_t^\infty r_v(z_f(\tau), u(\tau)) d\tau \right) \\ &= \int_t^\infty (z_f^T z_f + u^{*T} u^*) d\tau \end{aligned} \quad (6.67)$$

where Π is the set of admissible control inputs. The HJB equation associated with (6.67) is given by:

$$\begin{aligned} H_v\left(z_f, u^*, \frac{\partial I_v^*}{\partial z_f}\right) &= r_v(z_f, u^*) + \frac{\partial I_v^*}{\partial z_f} \dot{z}_f \\ &= z_f^T z_f + u^{*T} u^* + \frac{\partial I_v^*}{\partial z_f} \left(F + u - \dot{\alpha}_f^* \right) \end{aligned} \quad (6.68)$$

Following the same procedure as step 1, solving $\partial H_v\left(z_f, u^*, \frac{\partial I_v^*}{\partial z_f}\right) / \partial u^* = 0$ gives:

$$u^* = -\frac{1}{2} \frac{\partial I_v^*}{\partial z_f}(z_f) \quad (6.69)$$

To realize the control objective, the term $\frac{\partial I_v^*}{\partial z_f}$ is expressed as follows:

$$\frac{\partial I_v^*}{\partial z_f}(z_f) = 2C_v e_f + 2K_v z_f + 2F + I_v^0 \quad (6.70)$$

where $I_v^0 = -2C_v e_f - 2K_v z_f - 2F + \frac{\partial I_v^*}{\partial z_f}(z_f)$.

Inserting (6.70) into (6.69) gives:

$$u^* = -C_v e_f - K_v z_f - F - \frac{1}{2} I_v^0 \quad (6.71)$$

The unknown continuous functions F and I_v^0 can be approximated by RBFNNs as follows:

$$F = W_F^{*T} \Phi_F(X_F) + \epsilon_F(X_F) \quad (6.72)$$

$$I_v^0 = W_{I_v}^{*T} \Phi_{I_v}(X_{I_v}) + \epsilon_{I_v}(X_{I_v}) \quad (6.73)$$

where W_F^* and $W_{I_v}^*$ are the ideal weights, $\Phi_F(X_F)$ and $\Phi_{I_v}(X_{I_v})$ are the vectors basis functions, and $\epsilon_F(X_F)$ and $\epsilon_{I_v}(X_{I_v})$ are the RBFNN approximation errors.

Inserting (6.72) and (6.73) into (6.70) and (6.71), respectively gives:

$$\begin{aligned} \frac{\partial I_v^*(z_f)}{\partial z_f} &= 2C_v e_f + 2K_1 z_f + 2W_F^{*T} \Phi_F(X_F) \\ &\quad + W_{I_v}^{*T} \Phi_{I_v}(X_{I_v}) + \epsilon_v \end{aligned} \quad (6.74)$$

$$\begin{aligned} u^* &= -C_v e_f - K_v z_f - W_F^{*T} \Phi_F(X_F) \\ &\quad - \frac{1}{2} W_{I_v}^{*T} \Phi_{I_v}(X_{I_v}) + \frac{1}{2} \epsilon_v \end{aligned} \quad (6.75)$$

where $\epsilon_v = 2\epsilon_F(X_F) + \epsilon_{I_v}(X_{I_v})$.

The optimal control law (6.75) is unavailable since the weights W_F^* and $W_{I_v}^*$ are unknown. To obtain a usable optimized control, an RL architecture using the identifier, critic, and actor networks is constructed as:

$$\hat{F} = \hat{W}_F^T \Phi_F(X_F) \quad (6.76)$$

$$\begin{aligned} \frac{\partial \hat{I}_v(z_f)}{\partial z_f} &= 2C_v e_f + 2K_1 z_f + 2\hat{W}_F^{*T} \Phi_F(X_F) \\ &\quad + \hat{W}_{c_v}^{*T} \Phi_{I_v}(X_{I_v}) \end{aligned} \quad (6.77)$$

$$\begin{aligned}\hat{u}^* = & -C_v e_f - K_v z_f - \hat{W}_F^{*T} \Phi_F(X_F) \\ & - \frac{1}{2} \hat{W}_{a_v}^{*T} \Phi_{I_v}(X_{I_v})\end{aligned}\quad (6.78)$$

where $\hat{W}_F$, $\hat{W}_{c_v}^*$ and $\hat{W}_{a_v}^*$ are the estimates of the identifier, critic, and actor weights, respectively, $\hat{F}$ and $\frac{\partial \hat{I}_v(z_f)}{\partial z_f}$ are the estimates of F and $\frac{\partial I_v(z_f)}{\partial z_f}$, respectively.

The RBFNN weights of the identifier, critic, and actor are updated by the following update laws:

$$\dot{\hat{W}}_F = \Gamma_v \left(\Phi_F(X_F) z_f - \sigma_v \hat{W}_F \right) \quad (6.79)$$

$$\dot{\hat{W}}_{c_v} = -\gamma_{c_v} \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \quad (6.80)$$

$$\dot{\hat{W}}_{a_v} = -\Phi_{I_v}^T \Phi_{I_v} \left(\gamma_{a_v} \left(\hat{W}_{a_v} - \hat{W}_{c_v} \right) + \gamma_{c_v} \hat{W}_{c_v} \right) \quad (6.81)$$

where Γ_v is a positive-definite matrix, $\gamma_{a_v} > 0$ and $\gamma_{c_v} > 0$ are design constants, and $\sigma_v > 0$ is a small constant.

A Lyapunov candidate function is selected for the z_v -subsystem as:

$$\begin{aligned}L_2 = & L_1 + \frac{1}{2} z_f^T z_f + \frac{1}{2} \tilde{W}_F^T \Gamma_v^{-1} \tilde{W}_F \\ & + \frac{1}{2} \tilde{W}_{c_v}^T \tilde{W}_{c_v} + \frac{1}{2} \tilde{W}_{a_v}^T \tilde{W}_{a_v}\end{aligned}\quad (6.82)$$

where $\tilde{W}_F = \hat{W}_F - W_F^*$, $\tilde{W}_{a_v} = \hat{W}_{a_v} - W_{I_v}^*$, and $\tilde{W}_{c_v} = \hat{W}_{c_v} - W_{I_v}^*$ are the RBFNN weight errors for the identifier, actor, and critic networks, respectively.

Taking the time derivative of L_2 and using (6.79), (6.80), and (6.81), one has:

$$\begin{aligned}
\dot{L}_2 &= \dot{L}_1 + z_f^T \left(F + \hat{u}^* - \dot{\hat{\alpha}}_f^* \right) + \tilde{W}_F^T \left(\Phi_F^T z_f - \sigma_v \hat{W}_F \right) \\
&\quad - \gamma_{c_v} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
&\quad - \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \left(\gamma_{a_v} \left(\hat{W}_{a_v} - \hat{W}_{c_v} \right) + \gamma_{c_v} \hat{W}_{c_v} \right)
\end{aligned} \tag{6.83}$$

Inserting (6.76) and (6.78) into (6.83), one gets:

$$\begin{aligned}
\dot{L}_2 &= \dot{L}_1 + z_f \left[-C_v e_f - K_v z_f - \frac{1}{2} \hat{W}_{a_v}^T \Phi_{I_v} (X_{I_v}) \right. \\
&\quad \left. - \tilde{W}_F^T \Phi_F (X_F) + \epsilon_F - \dot{\hat{\alpha}}_f^* \right] + \tilde{W}_F^T \left(\Phi_F^T z_f - \sigma_v \hat{W}_F \right) \\
&\quad - \gamma_{c_v} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
&\quad - \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \left(\gamma_{a_v} \left(\hat{W}_{a_v} - \hat{W}_{c_v} \right) + \gamma_{c_v} \hat{W}_{c_v} \right)
\end{aligned} \tag{6.84}$$

Evaluating (6.84) yields:

$$\begin{aligned}
\dot{L}_2 &= \dot{L}_1 - z_f C_v e_f - z_f K_v z_f + z_f^T \epsilon_F - z_f^T \dot{\hat{\alpha}}_f^* \\
&\quad - \frac{1}{2} z_f^T \hat{W}_{a_v} \Phi_{I_v} - \sigma_v \tilde{W}_F^T \hat{W}_F - \gamma_{c_v} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
&\quad - \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \left(\gamma_{a_v} \left(\hat{W}_{a_v} - \hat{W}_{c_v} \right) + \gamma_{c_v} \hat{W}_{c_v} \right)
\end{aligned} \tag{6.85}$$

From Young's inequality, we get:

$$\begin{aligned}
z_f^T \epsilon_F &\leq \frac{1}{2} z_f^T z_f + \frac{1}{2} \epsilon_F^T \epsilon_F \\
-z_f^T \dot{\hat{\alpha}}_f^* &\leq \frac{1}{2} z_f^T z_f + \frac{1}{2} \dot{\hat{\alpha}}_f^{*T} \dot{\hat{\alpha}}_f^* \\
-z_f C_v e_f &\leq \frac{1}{2} z_f^T C_v z_f + \frac{1}{2} e_f^T C_v e_f \\
-\frac{1}{2} z_f^T \hat{W}_{a_v} \Phi_{I_v} &\leq \frac{1}{4} z_f^T z_f + \frac{1}{4} \hat{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v}
\end{aligned} \tag{6.86}$$

Substituting the Young's inequality (6.86) into (6.85) yields:

$$\begin{aligned}
\dot{L}_2 \leq & \dot{L}_1 + \frac{1}{2}e_f^T C_v e_f - z_f(K_v - C_v/2 - 1.75)z_f \\
& - \sigma_v \tilde{W}_F^T \hat{W}_F - \gamma_{c_v} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& - \gamma_{a_v} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} \\
& + (\gamma_{a_v} - \gamma_{a_c}) \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& + \frac{1}{4} \hat{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} + \frac{1}{2} \epsilon_F^T \epsilon_F + \frac{1}{2} \dot{\alpha}_f^{*T} \dot{\alpha}_f^*
\end{aligned} \tag{6.87}$$

Using $\tilde{W}_F = \hat{W}_F - W_F^*$, $\tilde{W}_{a_v} = \hat{W}_{a_v} - W_{a_v}^*$, and $\tilde{W}_{c_v} = \hat{W}_{c_v} - W_{c_v}^*$, the following facts are valid:

$$\tilde{W}_F^T \hat{W}_F = \frac{1}{2} \tilde{W}_F^T \tilde{W}_F + \frac{1}{2} \hat{W}_F^T \hat{W}_F - \frac{1}{2} W_F^{*T} W_F^* \tag{6.88}$$

$$\begin{aligned}
\tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} = & \frac{1}{2} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{c_v} + \frac{1}{2} \hat{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& - \frac{1}{2} W_{c_v}^{*T} \Phi_{I_v}^T \Phi_{I_v} W_{c_v}^*
\end{aligned} \tag{6.89}$$

$$\begin{aligned}
\tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} = & \frac{1}{2} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{a_v} + \frac{1}{2} \hat{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} \\
& - \frac{1}{2} W_{a_v}^{*T} \Phi_{I_v}^T \Phi_{I_v} W_{a_v}^*
\end{aligned} \tag{6.90}$$

Combining (6.88)-(6.90) with (6.87) yields:

$$\begin{aligned}
\dot{L}_2 \leq & \dot{L}_1 + \frac{1}{2} e_f^T C_v e_f - z_f (K_v - C_v/2 - 1.75) z_f \\
& - \frac{\sigma_v}{2} \tilde{W}_F^T \tilde{W}_F - \frac{\gamma_{c_v}}{2} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{c_v} \\
& - \frac{\gamma_{a_v}}{2} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{a_v} \\
& + (\gamma_{a_v} - \gamma_{a_c}) \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& - \frac{\gamma_{c_v}}{2} \hat{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& - \left(\frac{\gamma_{a_v}}{2} - \frac{1}{4} \right) \hat{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} \\
& + \left(\frac{\gamma_{a_v}}{2} + \frac{\gamma_{c_v}}{2} \right) W_{I_v}^{*T} \Phi_{I_v}^T \Phi_{I_v} W_{I_v}^* \\
& + \frac{\sigma_v}{2} W_F^{*T} W_F^* + \frac{1}{2} \epsilon_F^T \epsilon_F + \frac{1}{2} \dot{\alpha}_f^{*T} \dot{\alpha}_f^*
\end{aligned} \tag{6.91}$$

Based on Young's inequality, one gets.

$$\begin{aligned}
\tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} &= \frac{1}{2} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{a_v} \\
&+ \frac{1}{2} \hat{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v}
\end{aligned} \tag{6.92}$$

By utilizing (6.92), we get:

$$\begin{aligned}
\dot{L}_2 \leq & -e_f \hbar_f \Omega_{ff} (K_p - C_v/2 - 1.25) e_f \\
& - \frac{\gamma_{c_p}}{2} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{c_p} - \frac{\sigma_p}{2} \tilde{W}_\Upsilon^T \tilde{W}_\Upsilon \\
& - \frac{\gamma_{c_p}}{2} \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{a_p} - z_f^T (K_v - C_v/2 - 1.75) z_f \\
& - \frac{\gamma_{c_v}}{2} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{c_v} - \frac{\gamma_{c_v}}{2} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{a_v} \\
& - \frac{\sigma_v}{2} \tilde{W}_F^T \tilde{W}_F - \left(\gamma_{c_v} - \frac{\gamma_{a_v}}{2} \right) \hat{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{c_v} \\
& - \left(\frac{\gamma_{a_v}}{2} - \frac{1}{4} \right) \hat{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \hat{W}_{a_v} \\
& + \left(\frac{\gamma_{a_v}}{2} + \frac{\gamma_{c_v}}{2} \right) W_{I_v}^{*T} \Phi_{I_v}^T \Phi_{I_v} W_{I_v}^* \\
& + \frac{\sigma_v}{2} W_F^{*T} W_F^* + \frac{1}{2} \epsilon_F^T \epsilon_F + \frac{1}{2} \dot{\alpha}_f^{*T} \dot{\alpha}_f^* + \bar{\mathcal{D}}_1
\end{aligned} \tag{6.93}$$

The parameters C_v , K_v , γ_{c_v} , and γ_{a_v} are selected such that

$$\begin{aligned}
C_v &> 2.5, K_v > C_v, K_v - C_v > 3.5, \\
\gamma_{a_v} &> \frac{1}{2}, \gamma_{a_v} > \gamma_{c_v} > \frac{\gamma_{a_v}}{2}
\end{aligned}$$

Therefore, based on the selected parameters, (6.93) becomes:

$$\begin{aligned}
\dot{L}_2 \leq & -e_f^T (K_p - C_v/2 - 1.25) e_f - \frac{\gamma_{c_p}}{2} \tilde{W}_{c_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{c_p} \\
& - \frac{\sigma_p}{2} \tilde{W}_\Upsilon^T \tilde{W}_\Upsilon - \frac{\gamma_{c_p}}{2} \tilde{W}_{a_p}^T \Phi_{I_q}^T \Phi_{I_q} \tilde{W}_{a_p} \\
& - z_f^T (K_v - C_v/2 - 1.75) z_f - \frac{\sigma_v}{2} \tilde{W}_F^T \tilde{W}_F \\
& - \frac{\gamma_{c_v}}{2} \tilde{W}_{c_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{c_v} - \frac{\gamma_{c_v}}{2} \tilde{W}_{a_v}^T \Phi_{I_v}^T \Phi_{I_v} \tilde{W}_{a_v} + \mathcal{D}_2
\end{aligned} \tag{6.94}$$

where $\mathcal{D}_2 = \left(\frac{\gamma_{a_v}}{2} + \frac{\gamma_{c_v}}{2} \right) W_{I_v}^{*T} \Phi_{I_v}^T \Phi_{I_v} W_{I_v}^* + \frac{\sigma_v}{2} W_F^{*T} W_F^* + \frac{1}{2} \epsilon_F^T \epsilon_F + \frac{1}{2} \dot{\alpha}_f^{*T} \dot{\alpha}_f^* + \bar{\mathcal{D}}_1$, which

satisfies $\mathcal{D}_2 \leq \bar{\mathcal{D}}_2$, and $\bar{\mathcal{D}}_2 > 0$ is a constant.

The inequality (6.94) can be rewritten as:

$$\begin{aligned}
\dot{L}_2 \leq & -\lambda_q^{min} e_f^T e_f - \frac{\sigma_p}{2\lambda_{\Gamma_p^{-1}}^{max}} \tilde{W}_{\Upsilon}^T \Gamma_v^{-1} \tilde{W}_{\Upsilon} - \frac{\gamma_{c_p}}{2} \lambda_{I_q}^{min} \tilde{W}_{c_p}^T \tilde{W}_{c_p} \\
& - \frac{\gamma_{c_p}}{2} \lambda_{I_q}^{min} \tilde{W}_{a_p}^T \tilde{W}_{a_p} - \lambda_{I_v}^{min} z_f^T z_f - \frac{\sigma_v}{2\lambda_{\Gamma_v^{-1}}^{max}} \tilde{W}_F^T \Gamma_v^{-1} \tilde{W}_F \\
& - \frac{\gamma_{c_v}}{2} \lambda_{I_v}^{min} \tilde{W}_{c_v}^T \tilde{W}_{c_v} - \frac{\gamma_{c_v}}{2} \lambda_{I_v}^{min} \tilde{W}_{a_v}^T \tilde{W}_{a_v} + \bar{\mathcal{D}}_2
\end{aligned} \tag{6.95}$$

where λ_q^{min} is the minimum eigen value of $(K_p - C_v/2 - 1.25)$, λ_v^{min} is the minimum Eigen value of $(K_v - C_v/2 - 1.75)$, $\lambda_{I_q}^{min}$ is the minimum eigen value of $\Phi_{I_q}^T \hbar_f \Omega_{ff} \Phi_{I_q}$, $\lambda_{I_v}^{min}$ is the minimum eigen value of $\Phi_{I_v}^T \Phi_{I_v}$, $\lambda_{\Gamma_p^{-1}}^{max}$ is the maximum eigen value of Γ_p^{-1} , and $\lambda_{\Gamma_v^{-1}}^{max}$ is the maximum eigen value of Γ_v^{-1}

From (6.95), one can obtain:

$$\dot{L}_2 \leq -\mathcal{D}_3 L_2 + \bar{\mathcal{D}}_2 \tag{6.96}$$

where $\mathcal{D}_3 = \min\{2\lambda_q^{min}, (\sigma_p/\lambda_{\Gamma_p^{-1}}^{max}), (\gamma_{c_p}\lambda_{I_q}^{min}), (\gamma_{c_p}\lambda_{I_q}^{min}), 2\lambda_v^{min}, (\sigma_v/\lambda_{\Gamma_v^{-1}}^{max}), (\gamma_{c_v}\lambda_{I_v}^{min}), (\gamma_{c_v}\lambda_{I_v}^{min})\}$.

Theorem 6.1 *Consider the second-order nonlinear multi-agent system (6.6) with unknown nonlinear dynamics under deception attacks and actuator faults. By using the prescribed performance function (6.16), the error dynamics (6.24) and (6.25), the reinforcement learning-based optimized virtual controller (6.41) and the optimized overall controller (6.78) with identifier update laws (6.39),(6.79), critics update laws (6.48),(6.80), and actors update laws (6.49),(6.81), then the tracking errors in the*

closed-loop system are bounded and the leader-follower affine formation maneuvers are realized.

Proof.

Integrating both sides of (6.96) gives:

$$L_2(t) \leq e^{-\mathcal{D}_3 t} L_2(0) + \frac{\bar{\mathcal{D}}_2}{\mathcal{D}_3} \equiv \varrho \quad (6.97)$$

For all $t \geq T_s$, the inequality (6.97) shows that all the error signals e_f , z_f , $\tilde{W}_{a_p}$, $\tilde{W}_{a_v}$, $\tilde{W}_{c_p}$, $\tilde{W}_{c_v}$, $\tilde{W}_\Upsilon$, and $\tilde{W}_F$ are semi-globally bounded in finite settling time within a compact set defined by $\{L_2(t) : \|L_2(t)\| \leq \varrho\}$.

■

Remark 6.5 *It is obvious that increasing $\mathcal{D}_3$ or decreasing $\bar{\mathcal{D}}_2$ in (6.97) will aid in reducing $\bar{\mathcal{D}}_2/\mathcal{D}_3$ and subsequently makes the compact set smaller. This means that one may select the parameter $\mathcal{D}_3$ large enough to make e_f , z_f , $\tilde{W}_{a_p}$, $\tilde{W}_{a_v}$, $\tilde{W}_{c_p}$, $\tilde{W}_{c_v}$, $\tilde{W}_\Upsilon$, and $\tilde{W}_F$ sufficiently small.*

Remark 6.6 *Compared to the recent work of affine formation maneuver control of linear multi-agent systems [56], this article considers the issue of affine formation maneuver control of nonlinear multi-agent systems with preassigned performance. In addition, our proposed technique is able to counter deception attacks and actuator faults.*

Remark 6.7 *In [135], the authors used multiple adaptive laws to estimate the upper bounds of the attack signals Λ_{q_i} , Λ_{v_i} , $\Lambda_{q_i}\Lambda_{v_i}^{-1}$, $\dot{\Lambda}_{q_i}\Lambda_{q_i}^{-1}$, and $\dot{\Lambda}_{v_i}\Lambda_{v_i}^{-1}$. However,*

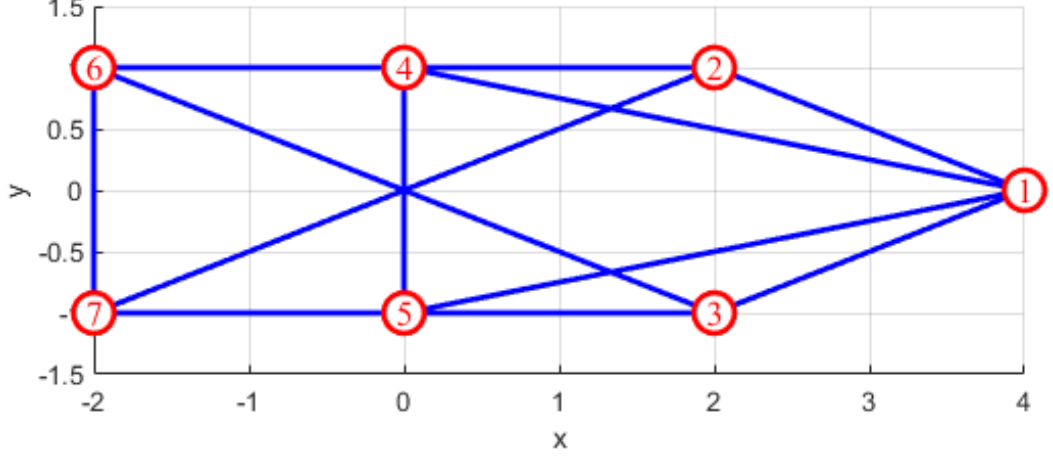


Figure 6.3: The leader-follower nominal formation topology.

the attack signals are usually time-varying and adaptive laws can only estimate constants. In this work, by transforming the multi-agent system to the form in (6.22) and (6.23), we used neural networks to approximate the lumped functions $\Upsilon(q_f, v_f, \Lambda_q, \Lambda_v)$ and $F(q_f, v_f, \Lambda_q, \Lambda_v, u)$ which have taken care of the time-varying attack signals and uncertain dynamics.

6.4 Simulation results

This section provides a numerical example to show the efficacy of the proposed method. The simulations are carried out on MATLAB/Simulink. The multi-agent system in this study consists of three leaders ($N_l = 3$) and four followers ($N_f = 4$) interacting over the nominal formation topology depicted in Fig. 6.3. The agents numbered 1, 2, and 3 are the leaders while the remaining agents numbered 4, 5, 6 and 7 are the followers. The nominal configurations of the seven agents are $q_1 = [4 \ 0]$, $q_2 = [2 \ 1]$, $q_3 = [2 \ -1]$, $q_4 = [0 \ 1]$, $q_5 = [0 \ -1]$, $q_6 = [-2 \ 1]$ and $q_7 = [-2 \ -1]$, and $q = [q_l^T, q_f^T]^T$,

where $q_l = [q_1^T, q_2^T, q_3^T]^T$ and $q_f = [q_4^T, q_5^T, q_6^T, q_7^T]^T$.

The target formation maneuvers of the leaders are described by the trajectories in (6.98). Various maneuvers of the leaders' nominal configurations can be obtained by manipulating $A(t)$ and $b(t)$.

$$q'_i(t) = A(t)q_i + b(t), \quad i = 1, \dots, N_l \quad (6.98)$$

As par Assumption 6.1, $q_l = q'_l$.

The motion of the followers is described by the following second-order nonlinear multi-agent systems.

$$\begin{aligned} \dot{q}_i &= v_i \\ \dot{v}_i &= \begin{bmatrix} h_{i1} \\ h_{i2} \end{bmatrix} + \begin{bmatrix} g_{i1} & 0 \\ 0 & g_{i2} \end{bmatrix} \begin{bmatrix} \check{u}_{i1} \\ \check{u}_{i2} \end{bmatrix} + \begin{bmatrix} \Phi_{i1}(t) \\ \Phi_{i2}(t) \end{bmatrix} \\ i &= N_l + 1, \dots, N_l + N_f \end{aligned} \quad (6.99)$$

with

$$\begin{bmatrix} \check{u}_{i1} \\ \check{u}_{i2} \end{bmatrix} = \begin{bmatrix} m_{i1} & 0 \\ 0 & m_{i2} \end{bmatrix} \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} + \begin{bmatrix} \delta_{i1} \\ \delta_{i2} \end{bmatrix} \quad (6.100)$$

where $h_{i1} = 2q_{i1}^2 + v_{i1}\cos(q_{i1})$, $h_{i2} = 2q_{i2}^2 + v_{i2}\cos(q_{i2})$, $g_{i1} = \sin(2v_{i1})$, $g_{i2} = \sin(2v_{i2})$, $\Phi_{i1}(t) = 0.12\sin(t)$, and $\Phi_{i2}(t) = 0.11 \cos(t)$, the loss of effectiveness faults are given by $m_{i1} = 0.9$, $m_{i2} = 0.8$, the bias faults are given by $\delta_{i1} = 0.1\sin(2t)\exp(-0.67t)$, $\delta_{i2} = 0.1\cos(2t) \exp(-0.02t)$.

The compromised sensor signals of the followers are given as:

$$\begin{cases} \bar{q}_i = \Lambda_{q_i} q_i \\ \bar{v}_i = \Lambda_{v_i} v_i \end{cases}, \quad i = N_l + 1, \dots, N_l + N_f \quad (6.101)$$

where $\Lambda_{q_i} = 1 - \cos^2(t)$ and $\Lambda_{v_i} = 1 + \sin(-3.46t)$.

The prescribed performance function is designed as:

$$\zeta_i(t) = \begin{cases} 2, & t = t_n \\ (2 - 0.2)\exp\left(\frac{-4(t-t_n)}{1}\right) + 0.2, & 0 < t - t_n < 1 \\ 0.2, & t - t_n \geq 1 \end{cases}$$

where $\zeta_{i(0)} = 2$, $\zeta_{i\infty} = 0.2$, $\kappa_i = 4$, $T_s = 1$ s, and $t_n = 10, 12, 17, 23, 29, 31$, and 33 s.

Whenever $t = t_n$, the prescribed performance function reset to $\zeta_i(t) = 2$.

The stress matrix is computed using the approach in [21] as follows:

$$\bar{\Omega} = \begin{bmatrix} 0.2878 & -0.2878 & -0.2878 & 0.1439 & 0.1439 & 0 & 0 \\ -0.2878 & 0.7194 & 0 & -0.5756 & 0 & 0 & 0.1439 \\ -0.2878 & 0 & 0.7194 & 0 & -0.5756 & 0.1439 & 0 \\ 0.1439 & -0.5756 & 0 & 0.7914 & -0.0719 & -0.2878 & 0 \\ 0.1439 & 0 & -0.5756 & -0.0719 & 0.7914 & 0 & -0.2878 \\ 0 & 0 & 0.1439 & -0.2878 & 0 & 0.2878 & -0.1439 \\ 0 & 0.1439 & 0 & 0 & -0.2878 & -0.1439 & 0.2878 \end{bmatrix}$$

A Gaussian function is chosen as the activation function of the radial basis function neural network. Each of the identifier, actor, and critic neural networks contains five nodes. The centers of the identifier, actor, and critic neural networks are equally spaced within $[-3 \ 3]$, $[-5 \ 5]$, and $[-5 \ 5]$, respectively. The widths of the identifier,

actor, and critic neural networks are 1.5, 2.5, and 2.5, respectively.

The initial weights of the identifier, actor, and critic neural networks are chosen as $\hat{W}_{i\Upsilon}(0) = 0.1_{5 \times 1}$ & $\hat{W}_{iF}(0) = 0.1_{5 \times 1}$, $\hat{W}_{ia_p}(0) = 0.2_{5 \times 1}$ & $\hat{W}_{ia_v}(0) = 0.2_{5 \times 1}$, and $\hat{W}_{ic_p}(0) = 0.1_{5 \times 1}$ & $\hat{W}_{ic_v}(0) = 0.1_{5 \times 1}$, respectively. The initial positions of the leaders and the followers are set as $q_1(0) = [4 \ 0]^T$, $q_2(0) = [2 \ 1]^T$ and $q_3(0) = [2 \ -1]^T$. $q_4(0) = [0.5 \ 1.5]^T$, $q_5(0) = [0.5 \ -0.5]^T$, $q_6(0) = [-1.75 \ 0.75]^T$, $q_7(0) = [-2 \ -2.4]^T$, respectively, with $q(0) = [q_1(0)^T, q_2(0)^T, \dots, q_7(0)^T]^T = [q_l(0)^T \ q_f(0)^T]^T$.

The parameters of the optimized virtual and real controllers are given as $K_q = 30_{8 \times 8}$, $C_v = 22_{8 \times 8}$, $K_v = 41_{8 \times 8}$, $\Gamma_\Upsilon = \Gamma_F = 2_{8 \times 8}$, $\sigma_p = \sigma_v = 0.3$, $\gamma_{a_p} = \gamma_{a_v} = 0.5$, $\gamma_{c_p} = \gamma_{c_v} = 0.8$

The simulation results are illustrated in. The reinforcement learning-based optimized virtual and actual control protocols of the followers are given in Fig. 6.5 and Fig. 6.6, respectively. These control laws protect the trajectories of the agents against the severity of deception attacks and actuator faults. The protected time-varying trajectories of the leader-follower multi-agent system are depicted in Fig. 6.6. The affine formation maneuvering tracking errors together with the proposed finite time preassigned performance function is presented in Fig. 6.7. It can be observed that whenever the leader-follower formation maneuver is transformed, the system experienced new transient states. With the aid of the proposed performance function, the new transient states of the tracking errors are contained within the required range, and they reached their steady states within the stipulated settling time. As shown in Fig. 6.8, affine formation maneuvers of the followers and the leaders are smooth and resilient to malicious deception attacks and actuator faults due to the utilization of

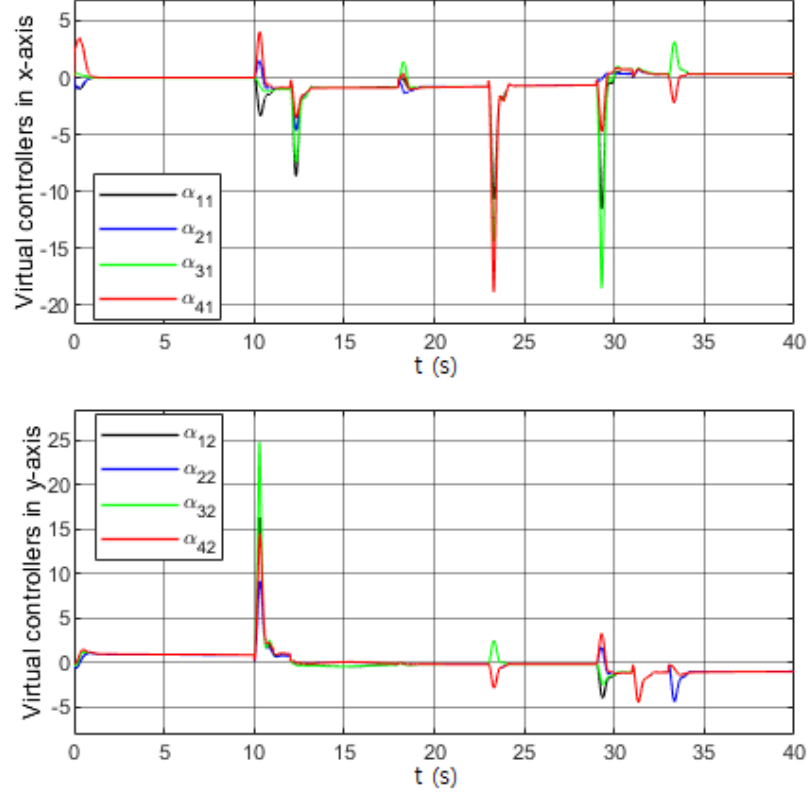


Figure 6.4: Optimized virtual control protocols of the four followers.

our control strategy.

To demonstrate the benefit of the proposed faulty-tolerant anti-attack affine formation control strategy, the affine formation maneuver control scheme developed in [163] is implemented on (6.6). The control approach in [163] was not designed to handle adversarial attacks or actuator faults. As a result, the impacts of the deception attacks and actuator faults manifested as shown in Fig. 6.9. The followers could not track the maneuvers of the leaders. Moreover, the followers collide with one another and could not maintain the target formation shape as shown in Fig. 6.10.

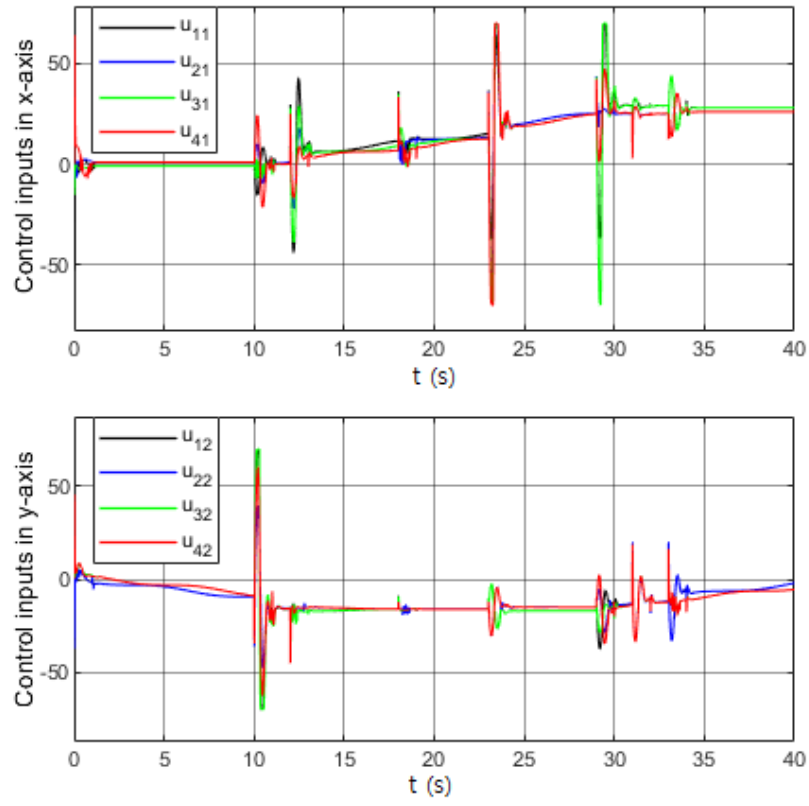


Figure 6.5: Optimized actual control protocols of the four followers.

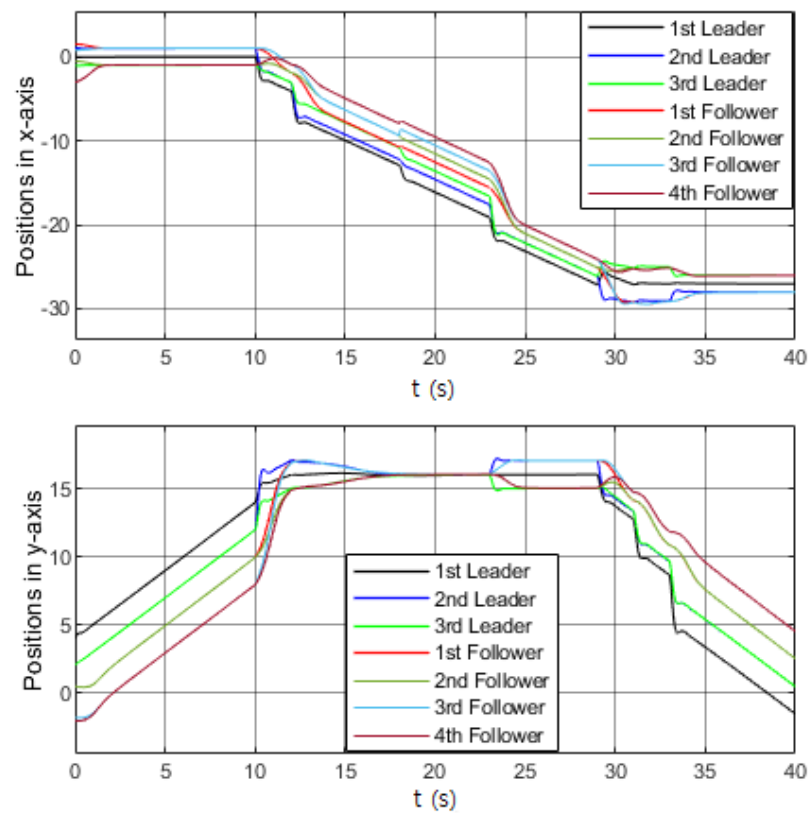


Figure 6.6: Time-varying trajectories of the followers and the leaders with control method.

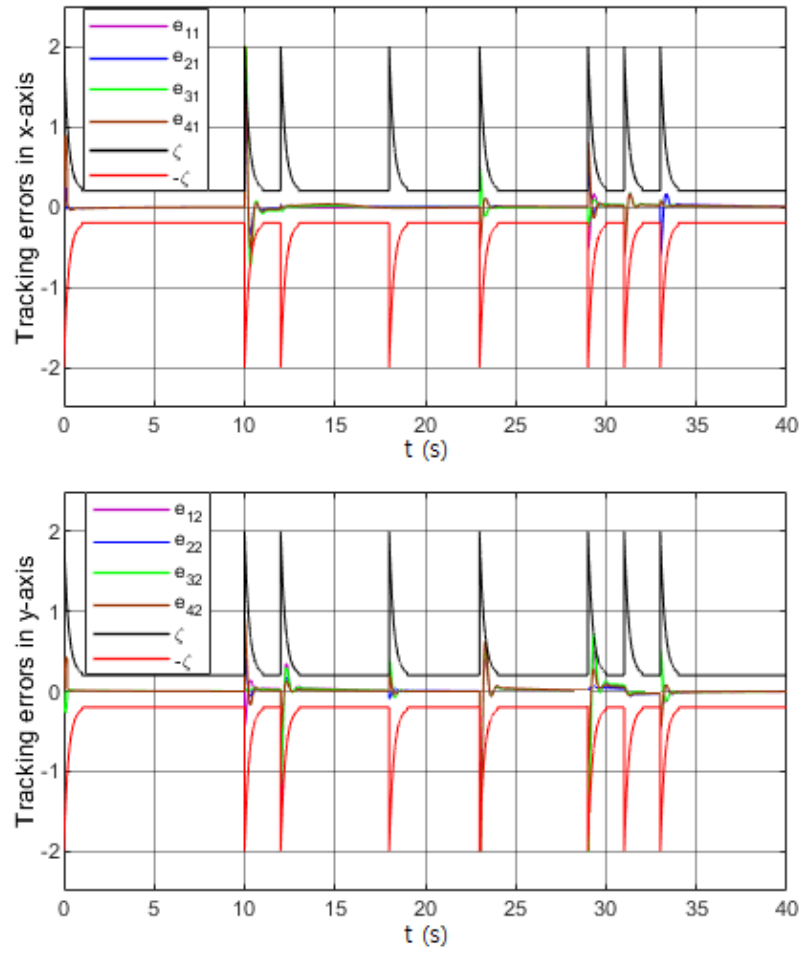


Figure 6.7: Prescribed leader-follower tracking errors.

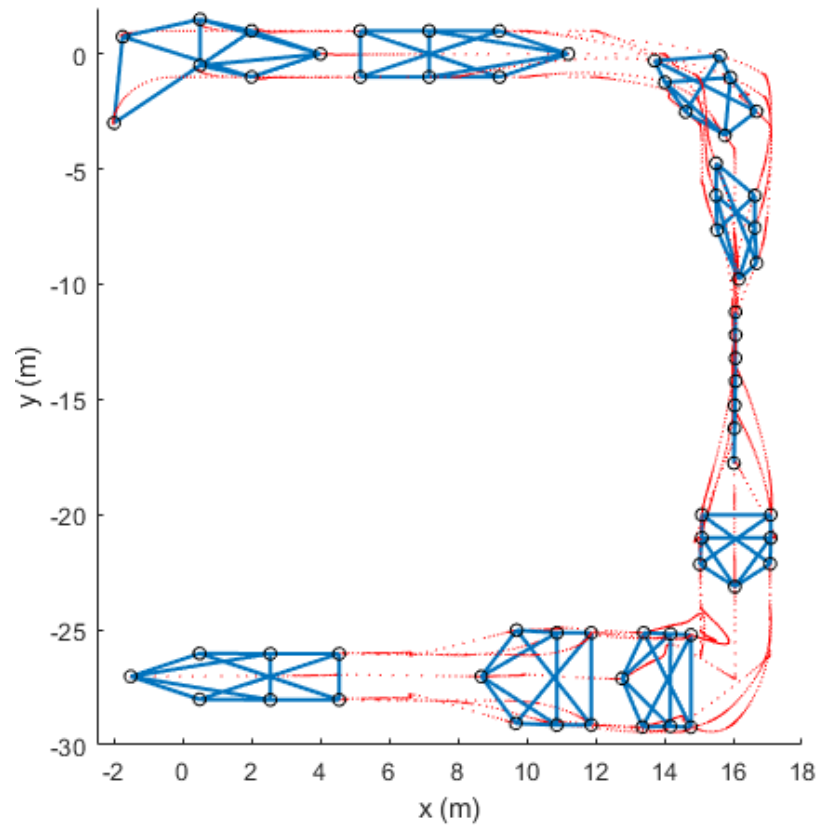


Figure 6.8: Various leader-follower affine formation maneuvering.

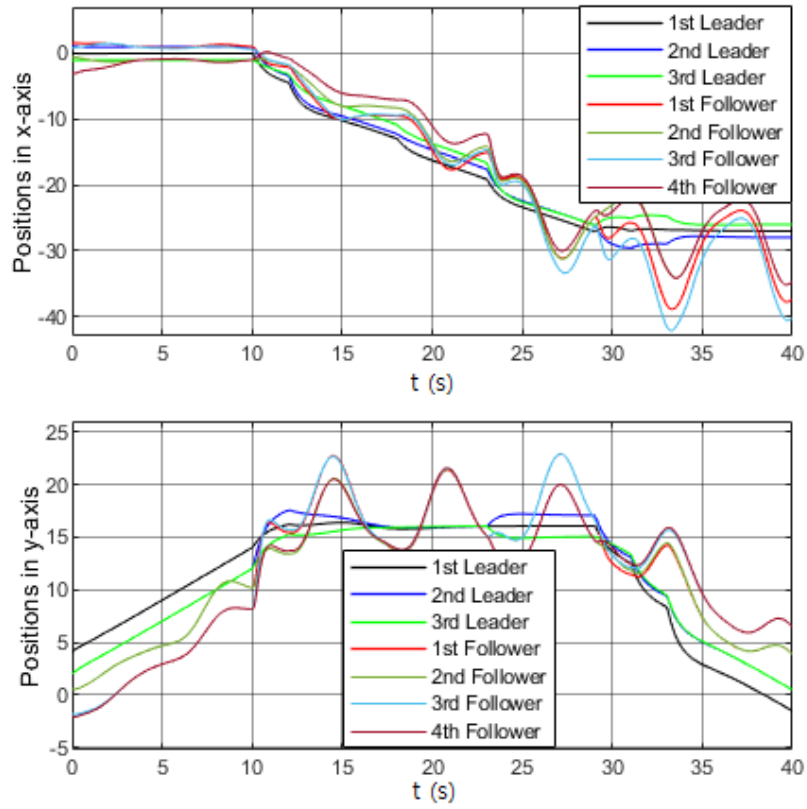


Figure 6.9: Time-varying trajectories of the followers and the leaders with control method in (Zhi et al. 2021).

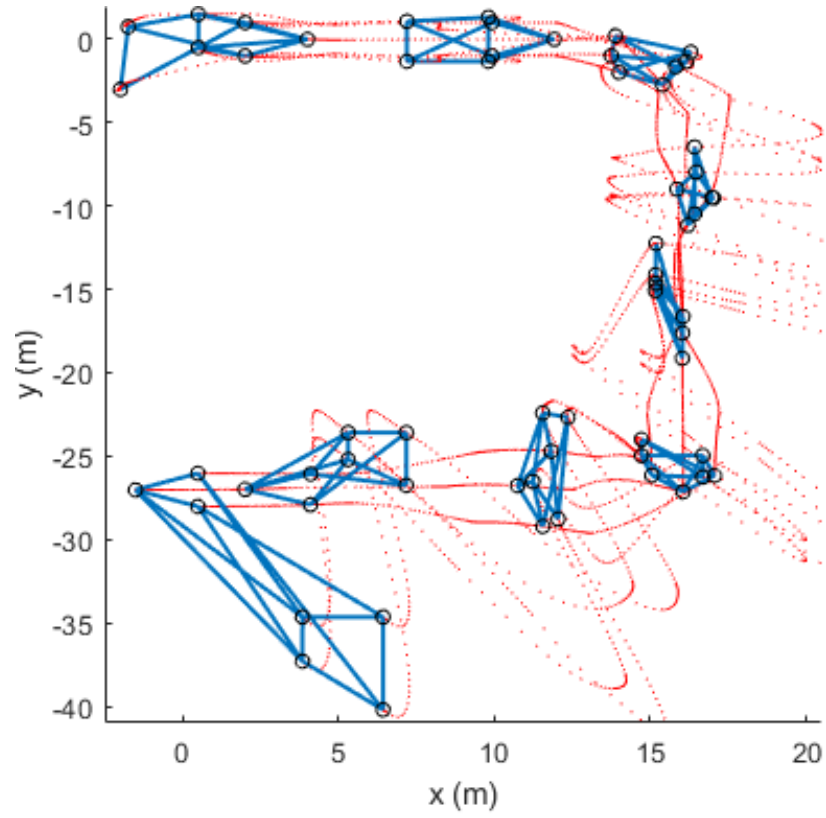


Figure 6.10: Various leader-follower affine formation maneuvering with control method in (Zhi et al. 2021).

CHAPTER 7

SAFETY-GUARANTEED SECURE REINFORCEMENT LEARNING-BASED AFFINE FORMATION MANEUVER CONTROL OF QUADROTORS IN ADVERSARIAL ENVIRONMENTS

In this chapter, an actor-critic learning scheme is studied for safe leader-follower affine formation maneuver control of networked quadrotors in the face of external disturbances, sensor deception attacks, and injection attacks on actuators. The followers aim to follow the formation maneuvers, such as scaling, shearing, translation, and rotation, decided by the leaders. Motivated by increasing safety and performance requirements during formation maneuvering, the dynamic states of the quadrotors are constrained within the prescribed safety region. To ensure that the safety constraints are not violated, a barrier Lyapunov function is put in place. Then, a distributed sliding mode control with actor-critic learning is formulated to facilitate accurate leader-follower affine formation maneuvers and reject malicious cyber-attack signals. The input gains that arise from attacks might corrupt the control direction. The Nussbaum gain function is coupled to the controller to tackle this problem. The actor system is responsible for estimating the uncertain dynamics together with the malicious attack signals while the critic network evaluates the control performance through

the estimated long-term performance index. The overall stability of the closed-loop system has been proven to be uniformly ultimately bounded using the Lyapunov stability theorem. Finally, the simulation results demonstrate the capability of the presented control method.

7.1 Introduction

In order to realize the required formation shape, a group of quadrotors requires information exchange among neighboring quadrotors through a wireless communication network. As a result, the quadrotors are exposed to cyber-attacks such as denial of service attacks, deception attacks, replay attacks, and false data injection attacks. Several efforts have been made to secure the communication networks of multiple quadrotors against harmful cyber-attacks. For instance, the authors in [142] proposed a singular perturbation approach to realize the formation control of quadrotors subjected to denial of service attacks. The authors in [175] focused on event-triggered formation control of a group of quadrotors subjected to denial of service attacks. In [122], the security formation control problem of multiple quadrotors facing deception attacks was investigated. By considering false data injection attacks in the actuators, the resilient control issue of a fleet of quadrotor UAVs has been solved in [176]. Nevertheless, At this moment, publications on multiple quadrotor systems under cyber-attacks are few. Moreover, no study has been reported on the formation control of quadrotors under multiple types of cyber-attacks.

Even though the above control schemes can guarantee the acquisition of the desired formation shape and stability of the quadrotors, the safety formation maneuver

within a constrained working space is overlooked. Therefore, the motion (vertical and horizontal displacements) of the quadrotors should be limited according to a defined safety range to improve their reliability, health, and life span. When the safety boundary is exceeded, the formation control performance of the system may be degraded, which can lead to operation failure. The barrier Lyapunov functions are commonly employed to prevent the violation of the dynamic states' constraints in control design [177]. In [177], barrier functions are utilized in the backstepping controller design to ensure safe formation acquisition of quadrotors while all constraints are satisfied. In [178], a barrier Lyapunov function is coupled with a prescribed performance function for time-varying leader-follower formation control of quadrotors. In [179], a barrier Lyapunov function together with a disturbance observer-based backstepping control strategy is employed to achieve leader-follower formation of quadrotors. Up to now, the formation control of quadrotors considering safety constraints is not being fully investigated.

The reinforcement learning algorithm, as a burning research area, is an effective intelligent approach capable of learning control policies without knowledge about the system's dynamics. The actor-critic architecture is widely deployed to solve reinforcement learning control problems [180]. Recently, the application of reinforcement learning in the formation control of multiple quadrotors is gradually attracting researchers [180], [181], [182], [183], [184]. The reinforcement learning method was used to learn the formation control protocol for a group of quadrotors with actuator faults in [180], [181]. The leader-follower formation tracking problem is addressed in [182] using optimal control technique through reinforcement learning. In [183], a distributed

leader-follower formation control based on reinforcement learning is developed for heterogeneous rotorcrafts. In [184], a reinforcement learning leader-follower optimal formation control of quadrotor swarm with switching communication graph is studied. Despite various publications on the reinforcement learning-based formation control of multiple quadrotors, to the best of our knowledge, the reinforcement learning-based distributed sliding mode formation control problem for multiple quadrotors with safety constraints and multiple types of cyber-attacks have not been studied yet.

The process of acquiring the target formation shape in [185]-[184] can be broadly classified as a bearing, distance, or relative-position approach. Each of the approaches has its own merits and shortcomings. For example, through the bearing approach, the quadrotors can accomplish scaling and translation formation maneuvers. On the other hand, the distance approach can realize formation translation and rotation maneuvers. Meanwhile, the formation shape described by the relative-position strategy can achieve a translation maneuver. In order to accomplish formation scaling, translation, and rotation simultaneously, affine formation maneuver control strategy based on stress matrix and affine transformation theory has been developed for multi-agent systems described by time-delayed linear dynamic equations [63], second-order integrator dynamic equations [21], [47], [186], triple-integrator dynamic equations [39], high-order linear dynamic equations [49], [56], and second-order nonlinear systems [163]. Designing formation maneuver control for multiple quadrotors using the stress matrix and affine transformation theory will give more freedom of maneuverability to the quadrotors. Therefore, affine formation maneuver control of networked quadrotors needs urgent attention.

In view of the foregoing discussion, this article aims to investigate the leader-follower affine formation maneuver control of networked quadrotors subjected to external disturbances, sensor deception attacks, and actuator injection attacks. By considering the safety and physical limitations of the quadrotors, their dynamic states are constrained to operate within a safe workspace. To prevent violation of the safety constraints, a barrier Lyapunov function is involved to guarantee the safety of the system. Then, the actor-critic learning-based distributed sliding mode control technique is designed to aid the leader-follower formation maneuvers of the quadrotors within the prescribed workspace while mitigating cyber-attacks. The key contributions of this chapter are summarized as follows:

1. By utilizing the properties of affine transformation and stress matrices, various collective affine formation maneuvers of the networked quadrotors such as scaling, translation, rotation, shearing, and collinearity can be accomplished contrary to [185]-[184]. Moreover, the quadrotors in this work have more freedom of maneuverability.
2. Safety-guaranteed leader-follower affine formation maneuver control problem of multiple quadrotors under sensor deception attacks and actuator injection attacks is addressed for the first time.
3. Compared with the works in [175], [122], [176] where the problem of input gains induced by the attack signals was not considered, this article addressed this problem by integrating the Nussbaum function into the proposed controller. As such, the proposed controller does not require any prior information about the

signs of the input gains induced by the attack signals.

4. Contrary to [180], [181], [182], [183], [184] where the learning-based controllers are in the form linear quadratic regulators, this paper proposed a distributed sliding mode control approach equipped with learning mechanism.

The remaining sections of this chapter are organized as follows: The preliminaries and the problem formulation are given in Section 7.2, the proposed control strategy is developed in Section 7.3, the simulation results are presented in Section 7.4.

7.2 Problem Formulation

7.2.1 Quadrotor Dynamic Model

The quadrotor is a six-degrees-of-freedom system with four control inputs. The detailed principle of the operation of quadrotors can be read in [187]. Assuming that the quadrotor framework is rigid and symmetric, and the center of gravity coincides with the body-frame origin, the dynamic equations of the displacement and rotation of the quadrotor are as follows [187]:

$$\ddot{\phi}_i = \frac{1}{I_x} \left[(I_y - I_z) \dot{\psi}_i \dot{\theta}_i - a_1 \dot{\phi}_i^2 - I_p \Omega \dot{\theta}_i + l u_{2i} \right] + \Delta_{\phi i} \quad (7.1)$$

$$\ddot{\theta}_i = \frac{1}{I_y} \left[(I_z - I_x) \dot{\psi}_i \dot{\phi}_i - a_2 \dot{\theta}_i^2 + I_p \Omega \dot{\phi}_i + l u_{3i} \right] + \Delta_{\theta i} \quad (7.2)$$

$$\ddot{\psi}_i = \frac{1}{I_z} \left[(I_x - I_y) \dot{\phi}_i \dot{\theta}_i - a_3 \dot{\psi}_i^2 + l u_{4i} \right] + \Delta_{\psi i} \quad (7.3)$$

$$\ddot{x}_i = -\frac{a_4 \dot{x}_i}{m} + (\cos \phi_i \sin \theta_i \cos \psi_i + \sin \phi_i \sin \psi_i) \frac{u_{1i}}{m} + \Delta_{xi} \quad (7.4)$$

$$\ddot{y}_i = -\frac{a_5 \dot{y}_i}{m} + (\cos \phi_i \sin \theta_i \sin \psi_i - \sin \phi_i \cos \psi_i) \frac{u_{1i}}{m} + \Delta_{yi} \quad (7.5)$$

$$\ddot{z}_i = -\frac{a_6 \dot{z}_i}{m} + (\cos \phi_i \cos \theta_i) \frac{u_{1i}}{m} - \frac{g}{m} + \Delta_{zi} \quad (7.6)$$

where x , y , and z denote the position of the quadrotors in the earth frame, θ , ψ , and ϕ represent the roll, yaw, and pitch angles, respectively, $\Delta_{\phi i}$, $\Delta_{\theta i}$, $\Delta_{\psi i}$, Δ_{xi} , Δ_{yi} , and Δ_{zi} stand for the unknown time-varying external disturbances, l is the length from the center of each actuator to the center of mass, I_p is the propeller inertia, I_x , I_y , I_z are the moment of inertia along the x, y, and z axes, respectively, g represents the gravitational acceleration, a_1 , a_2 , a_3 , a_4 , a_5 , and a_6 stand for the drag coefficients, u_{1i} , u_{2i} , u_{3i} , and u_{4i} denote the control inputs.

The three-degree of freedom translational equations of the N quadrotors are given by [188], [189], [142]

$$\ddot{x}_i = -\frac{a_4 \dot{x}_i}{m} + u_{xi} + \Delta_{xi} \quad (7.7)$$

$$\ddot{y}_i = -\frac{a_5 \dot{y}_i}{m} + u_{yi} + \Delta_{yi} \quad (7.8)$$

$$\ddot{z}_i = -\frac{a_6 \dot{z}_i}{m} + u_{zi} + \Delta_{zi} \quad (7.9)$$

where

$$\begin{aligned} u_{xi} &= (\cos \phi_i \sin \theta_i \cos \psi_i + \sin \phi_i \sin \psi_i) \frac{u_{1i}}{m} \\ u_{yi} &= (\cos \phi_i \sin \theta_i \sin \psi_i - \sin \phi_i \cos \psi_i) \frac{u_{1i}}{m} \\ u_{zi} &= (\cos \phi_i \cos \theta_i) \frac{u_{1i}}{m} - \frac{g}{m} \end{aligned}$$

Assuming that the quadrotors are flying at the same altitude, the translational equations of the quadrotors in two-dimensional space can be written as [190], [191]:

$$\ddot{q}_i = \bar{\Upsilon}_i(q_i, \dot{q}_i) + u_i + \Delta_i \quad (7.10)$$

where $q_i = [x_i \ y_i]^T \in \mathbb{R}^2$, $u_i = [u_{xi} \ u_{yi}]^T \in \mathbb{R}^2$, $\bar{\Upsilon}_i(q_i, \dot{q}_i) = [-\frac{a_4 \dot{x}_i}{m} \ -\frac{a_5 \dot{y}_i}{m}]^T \in \mathbb{R}^2$, $\Delta_i = [\Delta_{xi} \ \Delta_{yi}]^T \in \mathbb{R}^2$.

Remark 7.1 *The dynamic states of the quadrotors in (7.10) are constrained in the following compact sets:*

$$|q_i| = \begin{bmatrix} |x_i| \\ |y_i| \end{bmatrix} < \begin{bmatrix} k_{ci} \\ k_{ci} \end{bmatrix} \quad (7.11)$$

where $k_{ci} > 0$ is a constant. The compact set (7.11) specifies the region of safety operation of the quadrotors.

Assumption 7.1 *Let q_i^* be the target position of each quadrotor. It is assumed that q_i^* is bounded as*

$$|q_i^*| = \begin{bmatrix} |x_i^*| \\ |y_i^*| \end{bmatrix} < \begin{bmatrix} \Pi_i \\ \Pi_i \end{bmatrix} < \begin{bmatrix} k_{ci} \\ k_{ci} \end{bmatrix} \quad (7.12)$$

and its 1st and 2nd derivatives are bounded as follows:

$$|\dot{q}_i| = \begin{bmatrix} |\dot{x}_i| \\ |\dot{y}_i| \end{bmatrix} < \begin{bmatrix} \Pi'_i \\ \Pi'_i \end{bmatrix} < \begin{bmatrix} k_{ci} \\ k_{ci} \end{bmatrix} \quad (7.13)$$

$$|\ddot{q}_i| = \begin{bmatrix} |\ddot{x}_i| \\ |\ddot{y}_i| \end{bmatrix} < \begin{bmatrix} \Pi''_i \\ \Pi''_i \end{bmatrix} < \begin{bmatrix} k_{ci} \\ k_{ci} \end{bmatrix} \quad (7.14)$$

where $\Pi_i > 0$, $\Pi'_i > 0$ and $\Pi''_i > 0$ are constants,

The two-dimension translational equations of the quadrotors in the face of cyber-attacks are given as:

$$\begin{cases} \ddot{q}_i = \bar{\Upsilon}_i(q_i, \dot{q}_i) + \check{u}_i + \Delta_i \\ \check{q}_i = q_i + \delta_i(t, q_i) \\ \check{u}_i = \underline{h}_i u_i + \bar{h}_i(t, q_i, \dot{q}_i) \end{cases} \quad (7.15)$$

where $\check{q}_i \in \mathbb{R}^2$ is the vector of the compromised states, $\check{u}_i \in \mathbb{R}^2$ is the vector of the compromised control inputs, $\delta_i(t, q_i) \in \mathbb{R}^2$ is the vector of the deception attack signals, $\bar{h}_i = [\bar{h}_{xi} \ \bar{h}_{yi}]^T \in \mathbb{R}^2$ is the vector of time-varying injected attack signals, and $\underline{h}_i = \text{diag}\{\underline{h}_{xi} \ \underline{h}_{yi}\} \in \mathbb{R}^{2 \times 2}$

The deception attack signals can be described by state-dependent function as $\delta_i(t, q_i) = \kappa_i(t)q_i$ since they are mimicking the state variables, where $\kappa_i(t) \in \mathbb{R}^{2 \times 2}$

is a time-varying weight. Thus, the quadrotor states corrupted by the deception attacks are given as:

$$\check{q}_i = q_i + \kappa_i(t) \quad q_i = \varphi_i(t)q_i \quad (7.16)$$

where $\varphi_i(t) = (1 + \kappa_i(t))$

Remark 7.2 *As a result of the attack, the actual state variables q_i are no longer available. therefore, the contaminated state variables $\check{q}_i$ will be utilised for control design.*

From (7.16), we have:

$$\dot{\check{q}}_i = \dot{\varphi}_i q_i + \varphi_i \dot{q}_i \quad (7.17)$$

$$\ddot{\check{q}}_i = \ddot{\varphi}_i q_i + 2\dot{\varphi}_i \dot{q}_i + \varphi_i \ddot{q}_i \quad (7.18)$$

Noting that $q_i = \frac{\check{q}_i}{\varphi_i}$ and $\dot{q}_i = \frac{1}{\varphi_i} \left[\dot{\check{q}}_i - \dot{\varphi}_i \frac{\check{q}_i}{\varphi_i} \right]$ and (7.15), (7.18) becomes:

$$\begin{aligned} \ddot{\check{q}}_i &= \ddot{\varphi}_i \frac{\check{q}_i}{\varphi_i} + 2\frac{\dot{\varphi}_i}{\varphi_i} \left[\dot{\check{q}}_i - \dot{\varphi}_i \frac{\check{q}_i}{\varphi_i} \right] \\ &+ \varphi_i \bar{\Upsilon}_i + \varphi_i \Delta_i + \varphi_i \underline{h}_i u_i + \varphi_i \bar{h}_i \end{aligned} \quad (7.19)$$

The quadrotor systems under the cyber-attacks could be expressed as:

$$\ddot{\check{q}}_i = \Upsilon_i + g_i u_i \quad (7.20)$$

where $\Upsilon_i = \ddot{\varphi}_i \frac{\check{q}_i}{\varphi_i} + 2 \frac{\dot{\varphi}_i}{\varphi_i} \left[\dot{q}_i - \dot{\varphi}_i \frac{\check{q}_i}{\varphi_i} \right] + \varphi_i \bar{\Upsilon}_i + \varphi_i \Delta_i + \varphi_i \bar{h}_i$, and $g_i = \varphi_i \underline{h}_i$.

Remark 7.3 *The N quadrotors in (7.20) can be partitioned into two groups: the N_l leaders and $N_f = (N - N_l)$ followers.*

Let $\check{q}_l$ and $\check{q}_f$ be the vectors of the N_l leaders and N_f followers, respectively. Then:

$$\check{q}_l = [\check{q}_1^T \ \check{q}_2^T \ \dots \ \check{q}_{N_l}^T]^T \quad (7.21)$$

$$\check{q}_f = [\check{q}_{N_l+1}^T \ \check{q}_{N_l+2}^T \ \dots \ \check{q}_{N_l+N_f}^T]^T \quad (7.22)$$

The configuration of the N quadrotors which consists of their positions in two-dimensional space under cyber-attacks is thus:

$$\check{q} = [\check{q}_l^T \ \check{q}_f^T]^T = [\check{q}_1^T \ \check{q}_2^T \ \dots \ \check{q}_N^T]^T \in \mathbb{R}^{2N} \quad (7.23)$$

Remark 7.4 *It is essential to design a control law that can maintain the healthy states of the quadrotors within the safety constraints (7.11) even in the presence of cyber-attacks.*

Assumption 7.2 [192], [49] *It is assumed that the N_l leader quadrotors are already controlled to acquire the desired formation maneuvers. In this sense, the control procedure for the leader quadrotors will not be discussed.*

Definition 7.1 [193] *Given a continuous function $H(\varsigma) : \mathbb{R} \longrightarrow \mathbb{R}$ that satisfies:*

$$\begin{cases} \lim_{t \rightarrow \infty} \sup_t \frac{1}{t} \int_0^t N(\varsigma) d\varsigma = +\infty \\ \lim_{t \rightarrow \infty} \inf_t \frac{1}{t} \int_0^t N(\varsigma) d\varsigma = -\infty \end{cases} \quad (7.24)$$

then $H(\varsigma) : \mathbb{R} \longrightarrow \mathbb{R}$ is a Nussbaum function. Many functions satisfy the condition in (7.24), for instance, $\varsigma^2 \cos(\varsigma)$, $e^{\varsigma^2} \cos(1/2 \varsigma)$ and $\varsigma^2 \sin(\varsigma)$.

Lemma 7.1 Define the smooth functions $L(t)$ and $\varsigma(t)$ over the range $[0, t_f)$ with $L(t) \geq 0$, $H(\varsigma)$ is a smooth Nussbaum function. The following inequality holds:

$$L(t) \leq r_1 + e^{-r_2 t} \int_0^t [g(.)H(\varsigma) + 1] \dot{\varsigma} e^{r_2 t} dt \quad (7.25)$$

where $r_1 > 0$, $r_2 > 0$ are constants, $g(.)$ is a time-varying function, then, $L(t)$, $\varsigma(t)$, and $\int_0^t g(.)N(\varsigma)\dot{\varsigma}dt$ are bounded over $[0, t_f)$.

7.2.2 Radial Basis Function Neural Network

Radial basis function neural networks are commonly employed to approximate any smooth and continuous unknown term as:

$$\Upsilon = W^T \Theta(\chi) + \varepsilon(\chi) \quad (7.26)$$

where χ is the input vector, $W = [w_1 \ w_2 \ \dots \ w_k]^T \in \mathbb{R}^k$ is the weight vector, k number of nodes in the hidden layer, $\varepsilon(\chi)$ is the approximation error and $\|\varepsilon(\chi)\| \leq \bar{\varepsilon}$ with $\bar{\varepsilon} > 0$, $\Theta(\chi) = [\Theta_1(\chi) \ \Theta_2(\chi) \ \dots \ \Theta_k(\chi)]^T$ represent the basis function vector with entries $\Theta_i(\chi)$ given by

$$\Theta_i(\chi) = \exp \left(-\frac{\|(\chi - \bar{\mu}_i)\|^2}{\underline{\mu}_i^2} \right), \quad i = 1, 2, \dots, k \quad (7.27)$$

where $\bar{\mu}_i$ and $\underline{\mu}_i$ denote the center and width of the Gaussian function, respectively.

7.2.3 Barrier Function

A barrier function is widely employed to guarantee the safety of dynamic systems [194]. The barrier function tends to infinity when it reaches the barriers of the specified safety constraints, but never violates the constraints [195]. This property is utilized to design control laws that keep dynamic systems within the safety barriers.

Lemma 7.2 [195] *Suppose a Lyapunov candidate function $L(z)$ complying $L(z) \rightarrow \infty$ as $|z| \rightarrow c$, is deployed to ensure that the state-variables boundaries are not transgressed. Let the Lyapunov candidate function be chosen as follows:*

$$L(z) = \frac{1}{2} \log \frac{c^2}{c^2 - z^2} \quad (7.28)$$

where c is the constraint on z . It can be proven that L is positive definite and continuous within the set $|z| \leq c$. The control policy is developed to meet $\dot{L} \leq 0$.

Lemma 7.3 [196] *For any constant $c > 0$ and $z \in \mathbb{R}$ meeting $|z| \leq c$, one get*

$$\log \frac{c^2}{c^2 - z^2} < \frac{z^2}{c^2 - z^2} \quad (7.29)$$

7.3 Control Design

The main objective here is to develop an intelligent learning control method to ensure the safe formation maneuver of a system of quadrotors and shield them against harmful cyber-attacks. Considering Assumption 7.2 that the leader quadrotors are already controlled, only the control of the followers will be discussed. The distributed sliding

mode surface of the followers' affine formation maneuver tracking errors is designed as:

$$\begin{aligned}\zeta_i = & \sum_{j=1}^N s_{ij}(\dot{\check{q}}_i - \dot{\check{q}}_j) - \sum_{j=1}^N s_{ij}(\dot{\check{q}}_i(0) - \dot{\check{q}}_j(0)), \quad i \in \mathcal{V}_f \\ & + \Gamma_1 \sum_{j=1}^N s_{ij}(\check{q}_i - \check{q}_j) + \Gamma_2 \int_0^t s_{ij}(\Xi_i - \Xi_j)dt\end{aligned}\quad (7.30)$$

where $\Gamma_1 > 0$, $\Gamma_2 > 0$ are constant diagonal matrices, the function $\Xi_i(t)$ is design as:

$$\Xi_i(t) = \left[\sum_{j=1}^N s_{ij} \tanh^{\iota_1}(\check{q}_i - \check{q}_j) + \sum_{j=1}^N s_{ij} \tanh^{\iota_2}(\dot{\check{q}}_i - \dot{\check{q}}_j) \right] \quad (7.31)$$

where $1 < \iota_1 < 2$, and $\iota_2 > \iota_1$

The compact form of (7.30) can be expressed as:

$$\zeta_f = \bar{\Omega}_{ff}(\dot{\check{q}}_f - \dot{q}_f^*) - \dot{z}_f(0) + \Gamma_1 z_f + \Gamma_2 \int_0^t \Xi_f dt \quad (7.32)$$

where $z_f = \bar{\Omega}_{ff}(\check{q}_f - q_f^*)$, $q_f^* = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_2) \check{q}_l$ and $\dot{q}_f^* = -(\Omega_{ff}^{-1} \Omega_{fl} \otimes I_2) \dot{\check{q}}_l$ represent the target formation of the followers and its first-order derivative, respectively, and

$$\Xi_f = [\Xi_{N_l+1}^T \ \Xi_{N_l+1}^T \cdots \Xi_{N_l+N_f}^T]^T.$$

The main objective here is to achieve:

$$\lim_{t \rightarrow \infty} (\check{q}_f(t) - q_f^*(t)) = 0 \quad (7.33)$$

Differentiating (7.32) with respect to time gives:

$$\begin{aligned}\dot{\zeta}_f &= \bar{\Omega}_{ff} \left(\ddot{q}_f - \ddot{q}_f^* \right) + \Gamma_1 \dot{z}_f + \Gamma_2 \Xi_f \\ &= \bar{\Omega}_{ff} \left(\Upsilon_f + \varphi_f u_f - \ddot{q}_f^* + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f \right)\end{aligned}\quad (7.34)$$

where $\zeta_f = [\zeta_{N_l+1}^T \ \zeta_{N_l+1}^T \cdots \zeta_{N_l+N_f}^T]^T$, $z_f = [z_{N_l+1}^T \ z_{N_l+1}^T \cdots z_{N_l+N_f}^T]^T$, $\Upsilon_f = [\Upsilon_{N_l+1}^T \ \Upsilon_{N_l+2}^T \cdots \Upsilon_{N_l+N_f}^T]^T$, $\varphi_f = \text{diag}\{\varphi_{N_l+1} \ \varphi_{N_l+1} \cdots \varphi_{N_l+N_f}\}$, and $u = [u_{N_l+1}^T \ u_{N_l+2}^T \cdots u_{N_l+N_f}^T]^T$.

From (7.30), we know that $\zeta_i = [\zeta_{xi} \ \zeta_{yi}]^T$. Then, define a barrier Lyapunov function for the tracking error system (7.34) as follows:

$$L_1 = \bar{\Omega}_{ff}^{-1} \sum_{i=N_l+1}^{N_l+N_f} \frac{1}{2} \left(\frac{c_{xi}^2}{c_{xi}^2 - \zeta_{xi}^2} + \frac{c_{yi}^2}{c_{yi}^2 - \zeta_{yi}^2} \right) \quad (7.35)$$

where $c_{xi} > 0$ and $c_{yi} > 0$ are chosen to make sure that the states' constraints are not violated. From Remark 1 7.1 and Assumption 7.1, we set $c_{xi} = k_{ci} - \Pi_i$ and $c_{yi} = k_{ci} - \Pi_i$.

Differentiating L_1 with respect to time gives:

$$\begin{aligned}\dot{L}_1 &= \bar{\Omega}_{ff}^{-1} \sum_{i=N_l+1}^{N_l+N_f} \left(\frac{\zeta_{xi} \dot{\zeta}_{xi}}{c_{xi}^2 - \zeta_{xi}^2} + \frac{\zeta_{yi} \dot{\zeta}_{yi}}{c_{yi}^2 - \zeta_{yi}^2} \right) \\ &= \bar{\Omega}_{ff}^{-1} \sum_{i=N_l+1}^{N_l+N_f} \left(\begin{bmatrix} \frac{\zeta_{xi}}{c_{xi}^2 - \zeta_{xi}^2} & \frac{\zeta_{yi}}{c_{yi}^2 - \zeta_{yi}^2} \end{bmatrix} \begin{bmatrix} \dot{\zeta}_{xi} \\ \dot{\zeta}_{yi} \end{bmatrix} \right) \\ &= \bar{\Omega}_{ff}^{-1} \sum_{i=N_l+1}^{N_l+N_f} O_i^T \dot{\zeta}_i\end{aligned}\quad (7.36)$$

where $O_i = \begin{bmatrix} \frac{\zeta_{xi}}{c_{xi}^2 - \zeta_{xi}^2} & \frac{\zeta_{yi}}{c_{yi}^2 - \zeta_{yi}^2} \end{bmatrix}^T$.

Equation (7.36) can be rewritten as:

$$\begin{aligned}\dot{L}_1 &= O_f^T \bar{\Omega}_{ff}^{-1} \dot{\zeta}_f \\ &= O_f^T \left(\Upsilon_f + \varphi_f u_f - \ddot{q}_f^* + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f \right)\end{aligned}\quad (7.37)$$

where $O_f = [O_{N_l+1}^T \ O_{N_l+2}^T \ \dots \ O_{N_l+N_f}^T]^T$.

The control protocol for the quadrotors with loss of direction as a result of the actuator attack can be designed as:

$$\begin{aligned}\bar{u}_f &= \varphi_f^{-1} \begin{pmatrix} \hat{\Upsilon}_f + \Gamma_3 \zeta_f + \frac{1}{2} O_f + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f \\ + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f - \bar{\Omega}_{ff}^{-1} \ddot{q}_f^* \end{pmatrix} \\ u_f &= N(\varsigma_f) \bar{u}_f \\ \dot{\zeta}_f &= O_f^T \begin{pmatrix} \hat{\Upsilon}_f + \Gamma_3 \zeta_f + \frac{1}{2} O_f + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f \\ + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f - \bar{\Omega}_{ff}^{-1} \ddot{q}_f^* \end{pmatrix}\end{aligned}\quad (7.38)$$

where Γ_3 is a constant positive definite diagonal matrix.

Remark 7.5 *the control input $\bar{u}_f$ suffered from loss of direction as a result of the attack gain φ_f . The Nussbaum function $N(\varsigma_f)$ is employed to tackle this problem. The control input u_f is the corrected signal.*

Remark 7.6 *The control protocol (7.38) cannot be applied to the quadrotors directly because of the loss of control direction and the unknown vector $\hat{\Upsilon}_f$. Therefore, the design of actor-critic reinforcement learning scheme that will make (7.38) useful and meet the control objectives will be explained later.*

7.3.1 Critic design

The critic network is designed to assess the performance of the present control action and produce punishment/reward signals for adaptive learning. The critic network is designed using the neural network. Define a long-term cost function as:

$$J(t) = \int_t^{\infty} e^{-\frac{\tau-t}{\alpha}} \Lambda(\tau) d\tau \quad (7.39)$$

where $J = [J_{N_l+1}^T \ J_{N_l+1}^T \dots J_{N_l+N_f}^T]^T$, $\alpha > 0$ is a constant used to discount the future cost, $\Lambda(t)$ is the current cost function expressed as:

$$\Lambda(t) = \zeta_f^T Q \zeta_f + u_f^T R u_f \quad (7.40)$$

where Q and R are constant positive definite matrices.

The long-term cost function (7.39) contains future system information coupled with the compromised variables, so the solution is difficult to calculate. Thus, the critic neural network will be deployed to approximate J as

$$J = W_c^T \Theta_c(\chi_c) + \varepsilon_c \quad (7.41)$$

The critic neural network estimate of J is given by:

$$\hat{J} = \hat{W}_c^T \Theta_c(\chi_c) \quad (7.42)$$

where $\hat{W}_c$ is the weight estimate of W_c .

The continuous-time temporal difference error is obtained as:

$$\varrho_c(t) = \Lambda(t) - \frac{1}{\alpha} \hat{J}(t) + \dot{\hat{J}}(t) \quad (7.43)$$

For large α , i.e $\alpha \rightarrow \infty$, (7.43) simplifies to:

$$\varrho_c(t) = \Lambda(t) + \dot{\hat{J}}(t) = \Lambda(t) + \nabla \hat{J}(t) \dot{\chi}_c \quad (7.44)$$

where ∇ stands for the gradient operator.

Define the temporal difference error objective function as:

$$E_c = \frac{1}{2} \varrho_c^T \varrho_c \quad (7.45)$$

The weight update rule of the critic network is derived as follows:

$$\begin{aligned} \dot{W}_c &= -\sigma_c \frac{\partial E_c}{\partial W_c} = -\sigma_c \varrho_c(t) \frac{\partial \varrho_c}{\partial W_c} \\ &= -\sigma_c \varrho_c(t) \frac{\partial [\Lambda(t) - (1/\alpha) \hat{J}(t) + \dot{\hat{J}}(t)]}{\partial W_c} \\ &= -\sigma_c \varrho_c(t) \left[-\frac{1}{\alpha} \frac{\partial \hat{J}}{\partial W_c} + \frac{\partial}{\partial W_c} \left(\frac{\partial \hat{J}}{\partial \chi_c} \right) \right] \\ &= -\sigma_c (\Lambda(t) + W_c^T \beta) \beta \end{aligned} \quad (7.46)$$

where $\sigma_c > 0$ is the learning rate, $\beta = -\frac{\Theta_c}{\alpha} + \nabla \Theta_c \dot{\chi}_c$.

Design a Lyapunov function L_c for the critic network as:

$$L_c = \frac{1}{2} \tilde{W}_c^T \tilde{W}_c \quad (7.47)$$

Using (7.46) and (7.47), one gets:

$$\begin{aligned}
\dot{L}_c &= \tilde{W}_c^T \dot{\tilde{W}}_c = \tilde{W}_c^T \dot{W}_c \\
&= -\sigma_c \tilde{W}_c^T (\Lambda(t) + W_c^T \beta) \beta
\end{aligned} \tag{7.48}$$

At the end of the reinforcement learning, the weight update law (7.46) will minimize the objective function (7.45) such that $\varrho_c \rightarrow 0$ and $\frac{J}{\alpha} - \dot{J}$. It follows that:

$$\begin{aligned}
\Lambda(t) &= W_c^T \frac{\Theta_c}{\alpha} + \frac{\varepsilon_c}{\alpha} - \nabla I \dot{\chi}_c \\
&= W_c^T \frac{\Theta_c}{\alpha} + \frac{\varepsilon_c}{\alpha} - \nabla (W_c^T \Theta_c (\chi_c) + \varepsilon_c) \dot{\chi}_c \\
&= -W_c^T \beta + \epsilon_c
\end{aligned} \tag{7.49}$$

where $\epsilon_c = \varepsilon_c + \nabla \varepsilon_c \dot{\chi}_c$, $\|\varepsilon_c\| \leq \epsilon_{c,\max}$, with > 0 being the upper-bound.

Putting (7.49) into (7.48), one achieves:

$$\begin{aligned}
\dot{L}_c &= -\sigma_c \tilde{W}_c^T (\tilde{W}_c \Theta_c + \epsilon_c) \beta \\
&= -\sigma_c \beta^T \tilde{W}_c^T \tilde{W}_c - \sigma_c \tilde{W}_c^T \epsilon_c \beta \\
&\leq -\frac{\sigma_c \beta^T \beta}{2} \tilde{W}_c^T \tilde{W}_c + \frac{\sigma_c}{2} \epsilon_c^T \epsilon_c \\
&\leq -\frac{\sigma_c \beta^T \beta}{2} \tilde{W}_c^T \tilde{W}_c + \frac{\sigma_c}{2} \|\epsilon_{c,\max}\|^2
\end{aligned} \tag{7.50}$$

7.3.2 Actor design

The information about the lumped function Υ_f is unavailable due to the existence of cyber-attacks, unknown external disturbances, and uncertain dynamics. The actor neural network is deployed to approximate Υ_f as follows:

$$\Upsilon_f = W_a^T \Theta_c (\chi_a) + \varepsilon_a \tag{7.51}$$

The actor network's estimation of Υ_f is thus:

$$\hat{\Upsilon}_f = \hat{W}_a \Theta_a(\chi_a) \quad (7.52)$$

The current weight estimation error of the actor network is:

$$\tilde{\Upsilon}_f = \tilde{W}_a^T S_a(a_{\text{in}}) \quad (7.53)$$

where $\tilde{W}_a = W_A - \hat{W}_a$.

Let $J_d = 0$ be the desired cos-to-go in the actor network. The error between $\hat{J}$ and J_d is:

$$\tilde{J}_d = \hat{J} - J_d = \hat{J} \quad (7.54)$$

The total of the errors in the actor system is expressed as:

$$\varrho_a = \tilde{\Upsilon}_f + K_a \hat{J} \quad (7.55)$$

where K_a is a diagonal gain matrix. The weight update law of the actor network is:

$$\dot{\hat{W}}_a = -\sigma_a \frac{\partial E_a}{\partial \hat{W}_a} \quad (7.56)$$

where $\sigma_a > 0$ is the learning rate,

$$\begin{aligned} E_a &= \frac{1}{2} \varrho_a^T \varrho_a \\ \frac{\partial E_a}{\partial \hat{W}_a} &= \frac{\partial E_a}{\partial \varrho_a} \frac{\partial \varrho_a}{\partial \tilde{\Upsilon}_f} \frac{\partial \tilde{\Upsilon}_f}{\partial \hat{W}_a} = \left(\tilde{\Upsilon}_f + K_a \hat{J} \right) \Theta_a \end{aligned}$$

The fact that $\tilde{\Upsilon}_f$ is unknown, the weight update law is rewritten as:

$$\dot{\hat{W}}_a = -\sigma_a \left(\hat{W}_a^T \Theta_a (\chi_a) + K_a \hat{J} \right) \Theta_a \quad (7.57)$$

Select a candidate Lyapunov function for the actor network as follows:

$$L_a = \frac{1}{2} \tilde{W}_a^T \tilde{W}_a \quad (7.58)$$

The time-derivative of L_a yields:

$$\dot{L}_a = -\sigma_a \tilde{W}_a^T \Theta_a \left(\hat{W}_a^T \Theta_a (\chi_a) + K_a \hat{J} \right) \quad (7.59)$$

Considering that $\hat{J} = W_c^T \Theta_c (\chi_c) + \tilde{W}_c^T \Theta_c (\chi_c)$, we get:

$$\hat{J}^T \hat{J} \leq 2 (W_c^T \Theta_c)^T W_c \Theta_c + 2 \left(\tilde{W}_c^T \Theta_c \right)^T \tilde{W}_c^T \Theta_c \quad (7.60)$$

Inserting (7.60) into (7.59), one has:

$$\begin{aligned} \dot{L}_a &= -\sigma_a \tilde{W}_a^T \Theta_a \left(W_a^T \Theta_a (\chi_a) - \tilde{W}_a^T \Theta_a (\chi_a) + K_a \hat{J} \right) \\ &= -\sigma_a \tilde{W}_a^T \Theta_a W_a^T \Theta_a (\chi_a) + \sigma_a \left\| \tilde{W}_a \right\|^2 \left\| \Theta_a \right\|^2 \\ &\quad - \sigma_a \tilde{W}_a^T \Theta_a K_a \hat{J} \\ &\leq -\sigma_a \tilde{W}_a^T \Theta_a W_a^T \Theta_a (\chi_a) + \sigma_a \left\| \tilde{W}_a \right\|^2 \left\| \Theta_a \right\|^2 \\ &\quad - \frac{\sigma_a}{2} \left\| \tilde{W}_a \right\|^2 \left\| \Theta_a \right\|^2 + \frac{\sigma_a}{2} K_a^T K_a \left\| \hat{J} \right\|^2 \\ &\leq -\frac{\sigma_a}{2} \left\| \tilde{W}_a \right\|^2 \left\| \Theta_a \right\|^2 + \frac{\sigma_a}{2} \left\| W_a \right\|^2 \left\| \Theta_a \right\|^2 \\ &\quad - \frac{\sigma_a}{2} \left\| \tilde{W}_a \right\|^2 \left\| \Theta_a \right\|^2 \end{aligned} \quad (7.61)$$

$$\begin{aligned}
& + \sigma_a K_a^T K_a \left(\|W_c\|^2 \|\Theta_c\|^2 + \|\tilde{W}_c\|^2 \|\Theta_c\|^2 \right) \\
& = -\sigma_a \|\tilde{W}_a\|^2 \|\Theta_a\|^2 + \frac{\sigma_a}{2} \|W_a\|^2 \|\Theta_a\|^2 \\
& + \sigma_a K_a^T K_a \|\tilde{W}_c\|^2 \|\Theta_c\|^2 + \sigma_a K_a^T K_a \|W_c\|^2 \|\Theta_c\|^2
\end{aligned} \tag{7.62}$$

Using the actor estimation in (7.52), (7.38) becomes:

$$\begin{aligned}
\bar{u}_f &= \varphi_f^{-1} \begin{pmatrix} \hat{W}_a^T \Theta_a (\chi_a) + \Gamma_3 \zeta_f + \frac{1}{2} O_f + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f \\ + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f - \bar{\Omega}_{ff}^{-1} \ddot{q}_f^* \end{pmatrix} \\
u_f &= N(\zeta_f) \bar{u}_f \\
\dot{\zeta}_f &= O_f^T \begin{pmatrix} \hat{W}_a^T \Theta_a (\chi_a) + \Gamma_3 \zeta_f + \frac{1}{2} O_f \\ + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f - \bar{\Omega}_{ff}^{-1} \ddot{q}_f^* \end{pmatrix}
\end{aligned} \tag{7.63}$$

Equation (7.37) can be evaluated as follows:

$$\begin{aligned}
\dot{L}_1 &= O_f^T \dot{\zeta}_f = \dot{\zeta}_f - \dot{\zeta}_f + O_f^T \varphi_f u_f \\
&+ O_f^T \begin{pmatrix} W_a^T (\Theta_a \chi_a) + \varepsilon_a - \ddot{q}_f^* + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f \\ + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f \end{pmatrix} \\
&= \dot{\zeta}_f + \varphi_f N(\varsigma_f) \dot{\zeta}_f \\
&+ O_f^T \begin{pmatrix} W_a^T (\Theta_a \chi_a) + \varepsilon_a - \ddot{q}_f^* + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f \\ + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f \end{pmatrix} \\
&- O_f^T \begin{pmatrix} \hat{W}_a^T \Theta_a (\chi_a) + \Gamma_3 \zeta_f + \frac{1}{2} O_f \\ + \Gamma_1 \bar{\Omega}_{ff}^{-1} \dot{z}_f + \Gamma_2 \bar{\Omega}_{ff}^{-1} \Xi_f - \bar{\Omega}_{ff}^{-1} \ddot{q}_f^* \end{pmatrix} \\
&= -O_f^T \Gamma_3 \zeta_f - \frac{1}{2} O_f^T O_f + O_f^T \tilde{W}_a^T \Theta_a (\chi_a) b \\
&+ O_f^T \varepsilon_a + [\varphi_f N(\varsigma_f) + 1] \dot{\zeta}_f
\end{aligned} \tag{7.64}$$

Consider the following Young's inequalities:

$$O_f^T \varepsilon_a \leq \frac{\varpi O_f^T O_f}{2} + \frac{\varepsilon_a^T \varepsilon_a}{2\varpi} \tag{7.65}$$

$$-O_f^T \tilde{W}_a^T \Theta_a (\chi_a) \leq \frac{\varpi O_f^T O_f}{2} + \frac{\tilde{W}_a^T \Theta_a \Theta_a^T \tilde{W}_a}{2\varpi} \tag{7.66}$$

where $\varpi > 0$ is a constant

Using the inequalities above, (7.64) gives:

$$\begin{aligned}
\dot{L}_1 &\leq -O_f^T \Gamma_3 \zeta_f + O_f^T O_f \left(\varpi - \frac{1}{2} \right) + \frac{\tilde{W}_a^T \Theta_a \Theta_a^T \tilde{W}_a}{2\varpi} \\
&+ \frac{\varepsilon_a^T \varepsilon_a}{2\varpi} + [\varphi_f N(\varsigma_f) + 1] \dot{\zeta}_f
\end{aligned} \tag{7.67}$$

For $\varpi = \frac{1}{2}$, we have:

$$\begin{aligned} \dot{L}_1 \leq & -\lambda_{\min}(\Gamma_3) \sum_{i=N_l+1}^{N_l+N_f} \left(\frac{\zeta_{xi}^2}{c_{xi}^2 - \zeta_{xi}^2} + \frac{\zeta_{yi}^2}{c_{yi}^2 - \zeta_{yi}^2} \right) \\ & + \frac{\|\tilde{W}_a\|^2 \|\Theta_a\|^2}{2\varpi} + \frac{\|\varepsilon_a\|^2}{2\varpi} + [\varphi_f N(\varsigma_f) + 1] \dot{\varsigma}_f \end{aligned} \quad (7.68)$$

Theorem 7.1 *For the system of full states' constrained quadrotors, (7.20) facing multiple threats from deception and injection cyber-attacks, under Assumptions 7.1 and 7.2, by designing the safety-guaranteed secure reinforcement learning control law (7.63) together with critic learning (7.46) and actor learning (7.57), the scheme can ensure that (a) the negative effects of the cyber-attacks are diminished; (b) the state constraints are not violated and the quadrotors maintain operation within the safety boundary; (c) the tracking errors in the closed-loop system are bounded within a compact set; and (d) the leader-follower affine formation maneuver of the quadrotors is realized.*

Proof. Choose a Lyapunov function candidate as follows:

$$L = L_1 + L_a + L_c \quad (7.69)$$

I

where L_1 , L_c , and L_a are defined in (7.35), (7.47), and (7.58), respectively.

By differentiating L with respect to time, we have:

$$\begin{aligned}
\dot{L} \leq & -\lambda_{\min}(\Gamma_3) \sum_{i=N_l+1}^{N_l+N_f} \left(\frac{\zeta_{xi}^2}{c_{xi}^2 - \zeta_{xi}^2} + \frac{\zeta_{yi}^2}{c_{yi}^2 - \zeta_{yi}^2} \right) \\
& - \|\tilde{W}_a\|^2 \|\Theta_a\|^2 \left(\sigma_a - \frac{1}{2\varpi} \right) \\
& - \|\tilde{W}_c\|^2 \left(\frac{\sigma_c}{2} \|\beta\|^2 - \sigma_a \|K_a\|^2 \|\Theta_c\|^2 \right) \\
& + \sigma_a \|K_a\|^2 \|W_c\|^2 \|\Theta_c\|^2 + \frac{\sigma_a}{2} \|W_a\|^2 \|\Theta_a\|^2 \\
& + \frac{\sigma_c}{2} \|\epsilon_{c,\max}\|^2 + \frac{\|\varepsilon_a\|^2}{2\varpi} + [\varphi_f N(\varsigma_f) + 1] \dot{\varsigma}_f
\end{aligned} \tag{7.70}$$

Based on Lemma 7.3, one can gets:

$$\text{Log} \left[\frac{c_{xi}^2}{c_{xi}^2 - \zeta_{xi}^2} \right] \leq \frac{\zeta_{xi}^2}{c_{xi}^2 - \zeta_{xi}^2} \tag{7.71}$$

$$\text{Log} \left[\frac{c_{yi}^2}{c_{yi}^2 - \zeta_{yi}^2} \right] \leq \frac{\zeta_{yi}^2}{c_{yi}^2 - \zeta_{yi}^2} \tag{7.72}$$

Equation (7.70) can be simplified as:

$$\dot{L} \leq h_1 L + h_2 + [\varphi_f N(\varsigma_f) + 1] \dot{\varsigma}_f \tag{7.73}$$

where

$$\begin{aligned}
h_1 = \min & \left(2\lambda_{\min}(\Gamma_3), 2\|\Theta_a\|^2 \left(\sigma_a - \frac{1}{2\varpi} \right), \right. \\
& \left. 2\left(\frac{\sigma_c}{2} \|\beta\|^2 - \sigma_a \|K_a\|^2 \|\Theta_c\|^2 \right) \right) \\
h_2 = & \sigma_a \|K_a\|^2 \|W_c\|^2 \|\Theta_c\|^2 + \frac{\sigma_a}{2} \|W_a\|^2 \|\Theta_a\|^2 \\
& + \frac{\sigma_c}{2} \|\epsilon_{c,\max}\|^2 + \frac{\|\varepsilon_a\|^2}{2\varpi}
\end{aligned}$$

and

$$\begin{aligned} \left(\sigma_a - \frac{1}{2\varpi} \right) &> 0 \\ \left(\frac{\sigma_c}{2} \|\beta\|^2 - \sigma_a \|K_a\|^2 \|\Theta_c\|^2 \right) &> 0 \end{aligned}$$

Integrating (7.73) over the interval $[0 \ t]$ leads to:

$$L(t) \leq \frac{h_2}{h_1} + L(0)e^{-h_1 t} + e^{-h_1 t} \int_0^t [N(\varsigma) + 1] \dot{\varsigma} e^{h_1 \tau} d\tau \quad (7.74)$$

Selecting h_3 as the upper-bound of $e^{-h_1 t} \int_0^t [N(\varsigma) + 1] \dot{\varsigma} e^{h_1 \tau} d\tau$, we achieve:

$$L(t) \leq \frac{h_2}{h_1} + L(0)e^{-h_1 t} + h_3 \quad (7.75)$$

Based on the inequality above, the tracking error signals will in the end remain in the compact sets defined by:

$$\|\zeta_{xi}\| \leq c_{xi} \sqrt{1 - e^{-2\left(\frac{h_{2i}}{h_{1i}} + h_{3i} + L_{xi}(0)\right)}} \quad (7.76)$$

$$\|\zeta_{yi}\| \leq c_{yi} \sqrt{1 - e^{-2\left(\frac{h_{2i}}{h_{1i}} + h_{3i} + L_{yi}(0)\right)}} \quad (7.77)$$

$$\|\widetilde{W}_a\| \leq \sqrt{2 \left(\frac{h_2}{h_1} + h_3 + L(0) \right)} \quad (7.78)$$

$$\|\widetilde{W}_c\| \leq \sqrt{2 \left(\frac{h_2}{h_1} + h_3 + L(0) \right)} \quad (7.79)$$

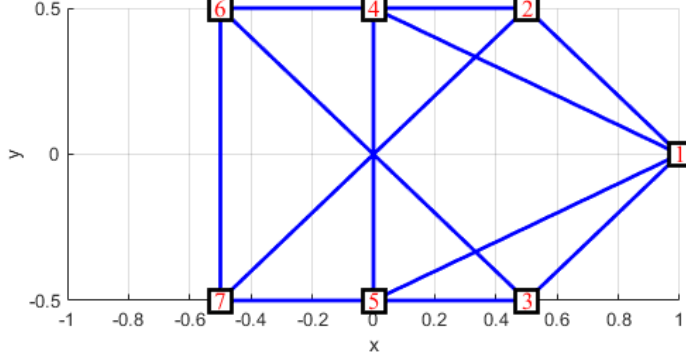


Figure 7.1: The nominal formation of the quadrotors.

7.4 Simulation results

In this section, simulations are given to highlight the validity of the developed control method.

The parameters of the quadrotor model are obtained from [187]. The nominal formation of the quadrotors is depicted in Fig. 7.1. The formation group comprises three leaders ($N_l = 3$) and four followers ($N_f = 4$). The quadrotors tagged (1, 2, & 3) and (4, 5, 6 & 7) represent the groups of leaders and followers, respectively, with corresponding nominal positions given as $\eta_1 = [1 \ 0]$, $\eta_2 = [0.5 \ 0.5]$, $\eta_3 = [0.5 \ -0.5]$, $\eta_4 = [0 \ 0.5]$, $\eta_5 = [0 \ -0.5]$, $\eta_6 = [-0.5 \ 0.5]$ and $\eta_7 = [-0.5 \ -0.5]$ and $\eta = [\eta_1^T, \eta_2^T, \dots, \eta_7^T]^T$. The stress matrix $\bar{\Omega}$ is computed according to the procedure detailed in [39]:

$$\bar{\Omega} = \begin{bmatrix} \bar{\Omega}_{ll} & \bar{\Omega}_{lf} \\ \bar{\Omega}_{fl} & \bar{\Omega}_{ff} \end{bmatrix} \quad (7.80)$$

where

$$\begin{aligned}
\bar{\Omega}_{fl} &= \begin{bmatrix} 0.1370 & -0.5482 & 0 \\ 0.1370 & 0 & -0.5482 \\ 0 & 0 & 0.1370 \\ 0 & 0.1370 & 0 \end{bmatrix} \\
\bar{\Omega}_{lf} &= \begin{bmatrix} 0.1370 & 0.1370 & 0 & 0 \\ -0.5482 & 0 & 0 & 0.1370 \\ 0 & -0.5482 & 0.1370 & 0 \end{bmatrix} \\
\bar{\Omega}_{ff} &= \begin{bmatrix} 0.7538 & -0.0685 & -0.2741 & 0 \\ -0.0685 & 0.7538 & 0 & -0.2741 \\ -0.2741 & 0 & 0.2741 & -0.1370 \\ 0 & -0.2741 & -0.1370 & 0.2741 \end{bmatrix} \\
\bar{\Omega}_{ll} &= \begin{bmatrix} 0.2742 & -0.2741 & -0.2741 \\ -0.2741 & 0.6853 & 0 \\ -0.2741 & 0 & 0.6853 \end{bmatrix},
\end{aligned}$$

The initial positions of the quadrotors are set as $q_1(0) = [1 \ 0]$, $q_2(0) = [0.5 \ 1]$, $q_3(0) = [0.5 \ -0.75]$, $q_4(0) = [0 \ 0.75]$, $q_5(0) = [0 \ -0.75]$, $q_6(0) = [-0.75 \ 1.5]$ and $q_7(0) = [-0.75 \ -1.25]$ and $q(0) = [q_1(0)^T, q_2(0)^T, \dots, q_7(0)^T]^T$. To enforce safety for the group of the quadrotors, the motion of each quadrotor is constrained in the safety region such that $|x_i| < k_{ci}$ and $|y_i| < k_{ci}$, with $k_{ci} = 15$. It follows that the tracking errors are also constrained as $|\zeta_{xi}| < c_{xi}$ and $|\zeta_{yi}| < c_{yi}$, with $c_{xi} = c_{yi} = 2$ to avoid violation of the safety region.

The sensor deception attacks are modeled as $\varphi_i = (1 + \cos(t))$, the actuator injection attacks are modeled as $m_i = \sin(q_i(t)\dot{q}_i(t))$.

The controller gains are chosen as $\Gamma_1 = \text{diag}\{15, 15, \dots, 15\}$, $\Gamma_2 = \text{diag}\{2, 2, \dots, 2\}$, $\Gamma_3 = \text{diag}\{10, 10, \dots, 10\}$, $\iota_1 = 1.5$, and $\iota_2 = 2.1$. The parameters in the cost function $R = \text{diag}\{5, 5, \dots, 5\}$, $Q = \text{diag}\{10, 10, \dots, 10\}$ and the discount factor $\alpha = 0.2$. The reinforcement learning parameters are given as: 5 nodes are considered in both actor and critic neural networks; the learning rates of the actor and critic neural networks are $\sigma_a = 1.5$ and $\sigma_c = 0.01$, respectively; the centers of the Gaussian functions of both actor and critic networks are selected between -0.5 and 0.5 , while and the width of both functions are set as 0.25 ; the initial weights of the actor and critic networks are selected as $\hat{W}_{ai}(0) = \hat{W}_{ci}(0) = [0.5, 0.5, \dots, 0.5]^T$.

Through the developed control protocols, the simulation results can be obtained as shown in Figs. 7.2-7.7. Fig. 7.2 illustrates the leader-follower time-varying trajectories of the quadrotors under sensor deception attacks and actuator injection attacks. It can be seen that the proposed learning-based controller is able to keep the quadrotors within the safety region despite the cyber-attacks. Moreover, the leader-follower tracking errors did not violate the prescribed constraints as presented in Fig. 7.3. The responses of the actor-critic learning-based control protocols that aid the realization of the aforementioned control performance are depicted in Fig. 7.4. The evolution of the norm-2 of the weight vectors of the actor and critic networks are depicted in Fig. 7.6 and Fig. 7.7. The affine formation maneuvers of the leader-follower multiple-quadrotor system are shown in Fig. 7.7.

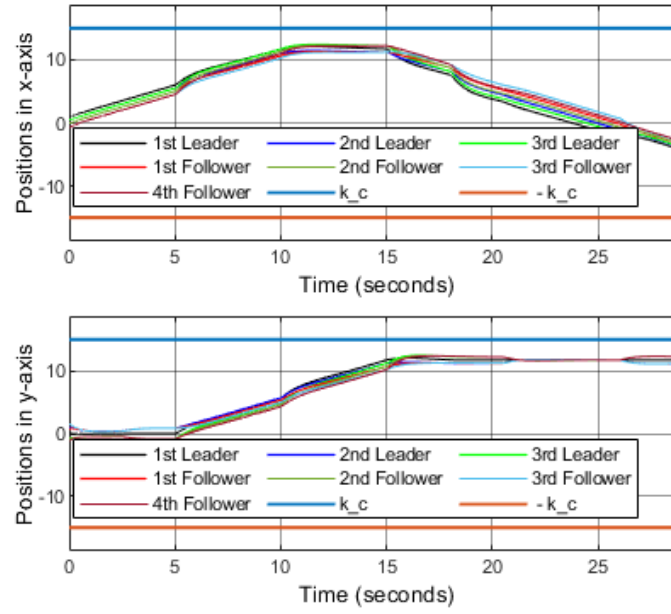


Figure 7.2: Leader-follower Time-varying positions trajectories.

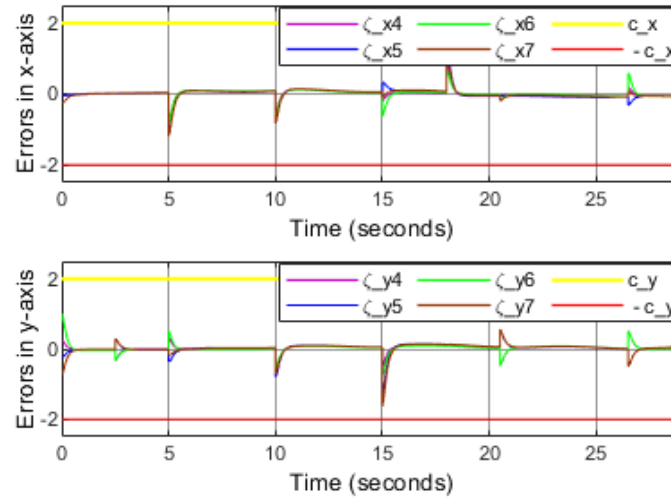


Figure 7.3: Tracking errors of the followers.

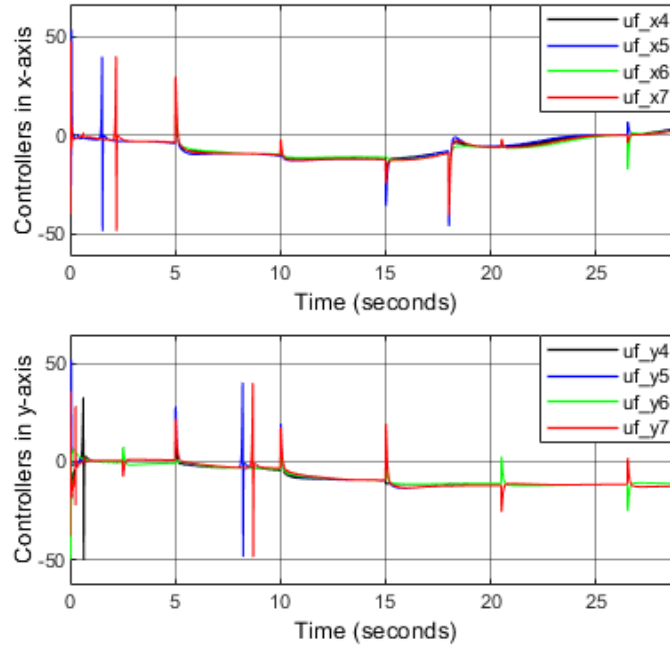


Figure 7.4: Actor-critic learning control protocols of the followers.

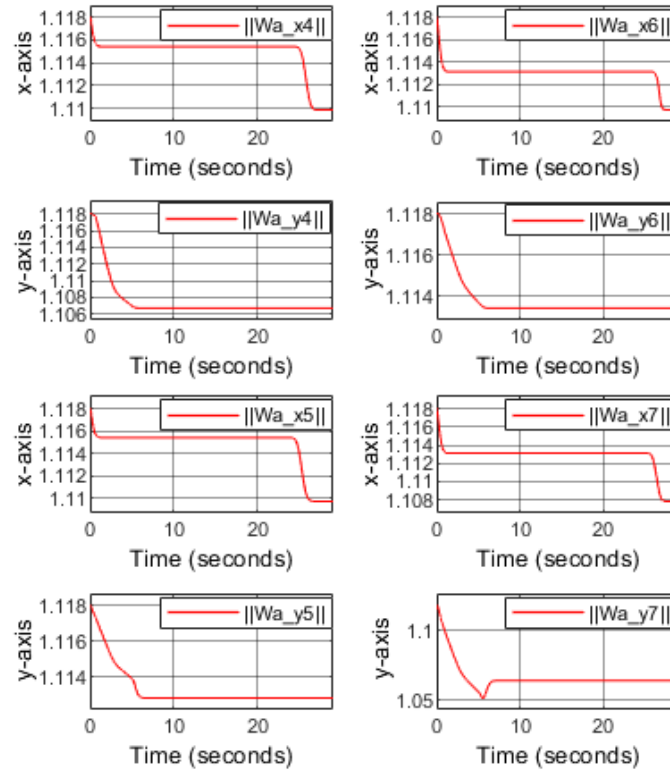


Figure 7.5: Norm-2 of the actor weights.

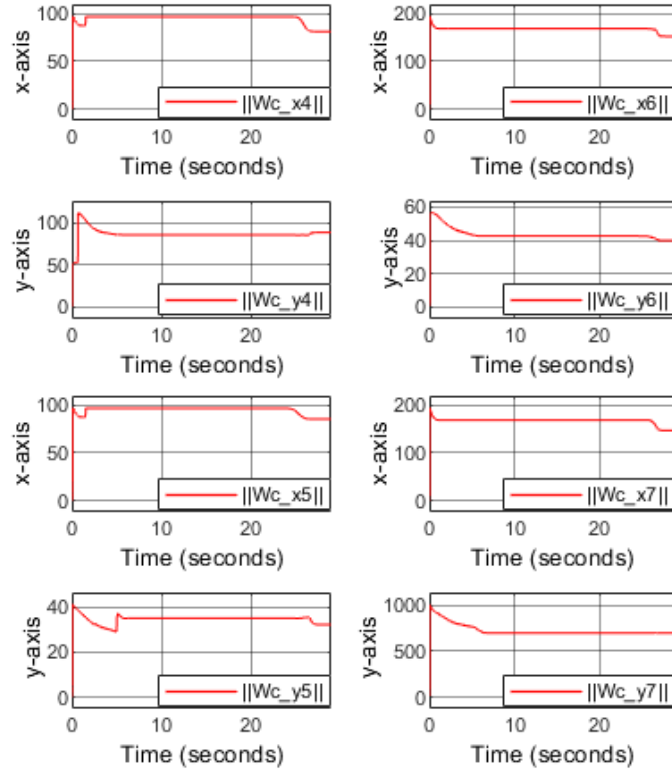


Figure 7.6: Norm-2 of the critic weights.

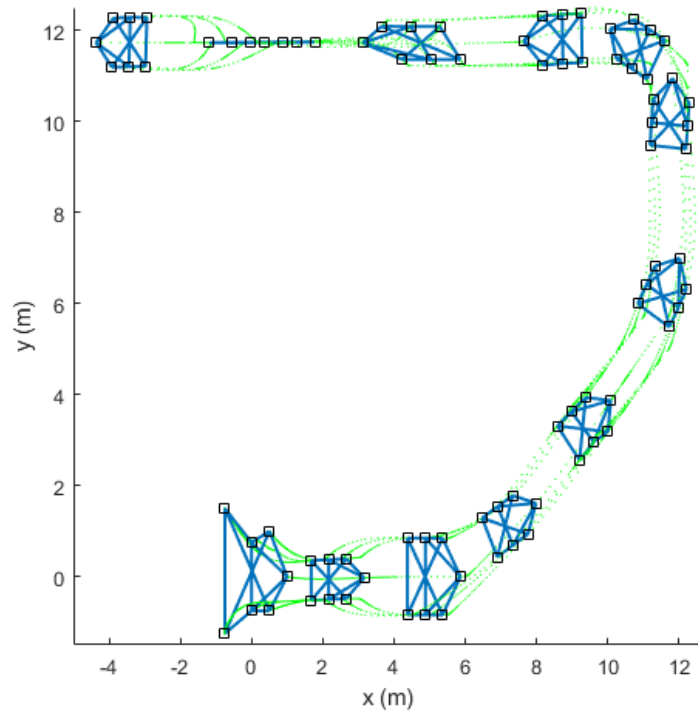


Figure 7.7: Leader-follower affine formation maneuvers.

CHAPTER 8

NEURAL NETWORK-BASED CONTROL FOR AFFINE FORMATION MANEUVER OF MULTI-AGENT SYSTEMS WITH EXTERNAL DISTURBANCES

This chapter proposes a distributed control law with a neural network to achieve leader-follower affine formation maneuver (AFM) of multi-agent systems (MASs) subjected to adversarial disturbances. The leaders determine the desired collective formation maneuvering such as scaling, shearing, translation, and rotation. The distributed controller ensures that the followers are tracking the maneuvers of the leaders at all times despite the disturbances. The disturbances are approximated and compensated with the aid of neural networks. A Lyapunov candidate function is used to highlight that the closed-loop system is uniformly ultimately bounded. Finally, in order to validate the proposed control method, it is applied to a multi-agent system of quadrotors. The simulation results have shown that the proposed control protocol is able to maintain the collective formation maneuver.

8.1 Introduction

In this chapter, an artificial neural network-based distributed control approach is proposed for a second-order MAS in the presence of adversarial disturbances. Unlike [68], [197] where adaptive laws were used to estimate uncertain parameters or constant disturbances, a neural network is used to estimate and mitigate adversarial disturbances to guarantee resilient leader-follower affine formation maneuvering. The bounded stability of the closed-loop system is guaranteed using a Lyapunov candidate function.

8.2 Problem formulation

This paper considers a multi-agent system consisting of N agents such that the first N_l agents represent the leaders while the remaining $N_f = N - N_l$ agents represent the followers. The dynamic equations of each follower are given by:

$$\begin{cases} \dot{q}_i = u_i, \\ \dot{u}_i = v_i + d_i, \quad i \in \mathcal{V}_f, \end{cases} \quad (8.1)$$

where $q_i \in \mathbb{R}^n$ is the position of each agent, $u_i \in \mathbb{R}^n$ is the velocity of each agent, $v_i \in \mathbb{R}^n$ is the control protocol of each agent, and $d_i \in \mathbb{R}^n$ represents the adversarial disturbances.

For the N agents, we can obtain the following vectors: $w = [q_1^T, \dots, q_N^T]^T = [q_l^T, q_f^T]^T \in \mathbb{R}^{nN}$ and $u = [u_1^T, \dots, u_N^T]^T = [u_l^T, u_f^T]^T \in \mathbb{R}^{nN}$, respectively.

Assuming that the desired position trajectories of the leaders $q_l^* = [q_1^{*T}, \dots, q_{N_l}^{*T}]^T$ are known in advance, then, based on definition 1, the desired position trajectories of

the followers $q_f^* = [q_{N_l+1}^{*T}, \dots, q_N^{*T}]^T$ can be calculated as:

$$q_f^* = -\bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} q_l^* \quad (8.2)$$

The tracking errors of the positions and velocities of the followers are thus:

$$\varepsilon_{q_f} = q_f - q_f^* = q_f + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} q_l^* \quad (8.3)$$

$$\varepsilon_{u_f} = u_f - u_f^* = u_f + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} u_l^* \quad (8.4)$$

The main objective of this work is to design v such that $\varepsilon_{q_f}(t), \varepsilon_{u_f}(t) \rightarrow 0$ as $t \rightarrow \infty$.

8.2.1 Neural network

Due to the universal approximation property of neural networks, they are widely utilized to approximate any continuous and smooth nonlinear function $(\Theta(X))$ as follows[198]:

$$\Theta(X) = W^T \Pi(X) + \delta(X) \quad (8.5)$$

where W is the weight vector of the neural networks, $\Pi(X)$ is the vector of the basis functions, and $\delta(X)$ is the approximation error satisfying $\|\delta(X)\| \leq \bar{\delta}$, where $\bar{\delta}$ is the upper-bound. In this article, the neural networks will be used to estimate and suppress the adversarial disturbances hampering the performance of the MAS.

8.3 Control design

In this section, a leader-follower distributed control protocol is designed to achieve $\varepsilon_{q_f}(t), \varepsilon_{u_f}(t) \rightarrow 0$ despite the adversarial disturbances. The adversarial disturbances are approximated and compensated using neural networks. This will guarantee the convergence of the followers to the desired formation maneuvers established by the leaders. The distributed control protocol is designed as follows:

$$v_i = -\frac{1}{\alpha_i} \sum_{j \in \mathcal{N}_i} \omega_{ij} [k_q (q_i - q_j) + k_u (u_i - u_j) - \dot{u}_j] - \hat{d}_i \quad (8.6)$$

$$\hat{d}_i = \hat{W}_i^T \Phi_i(X) \quad (8.7)$$

$$\dot{\hat{W}}_i = \sigma (\alpha_i \otimes I_n) \Phi_i(X) \varepsilon_{u_{f,i}} \quad (8.8)$$

where $\alpha_i = \sum_{j \in \mathcal{N}_i} \omega_{ij}$, $\sigma > 0$, $k_q > 0$ and $k_u > 0$ are design constants, $\hat{d}$ is the estimate of d , $\hat{W}_i$ is the estimate of W_i .

The disturbance approximation error $\tilde{d}_i$ is given by:

$$\begin{aligned} \tilde{d}_i &= d_i - \hat{d}_i = W_i^T \Phi_i(X) + \delta_i(X) - \hat{W}_i^T \Phi_i(X) \\ &= \tilde{W}_i^T \Phi_i(X) + \delta_i(X) \end{aligned} \quad (8.9)$$

where $\tilde{W}_i = W_i - \hat{W}_i$ is the weight estimation error. Substituting (8.6), (8.7), (8.8) and (8.9) into (8.1) yields:

$$\begin{aligned}
\dot{q}_i &= u_i \\
\dot{u}_i &= -\frac{1}{\alpha_i} \sum_{j \in \mathcal{N}_i} \omega_{ij} [k_q (q_i - q_j) + k_u (u_i - u_j) - \dot{u}_j] \\
&\quad + \widetilde{W}_i^T \Phi_i(X) + \delta_i(X)
\end{aligned} \tag{8.10}$$

Multiplying $\alpha_i = \sum_{j \in \mathcal{N}_i} \omega_{ij}$ on both sides of (8.10) gives:

$$\begin{aligned}
\sum_{j \in \mathcal{N}_i} \omega_{ij} (\dot{u}_i - \dot{u}_j) &= \sum_{j \in \mathcal{N}_i} \omega_{ij} [-k_q (q_i - q_j) - k_u (u_i - u_j)] \\
&\quad + \alpha_i \widetilde{W}_i^T \Phi_i(X) + \alpha_i \delta_i(X)
\end{aligned} \tag{8.11}$$

Equation (8.11) can be transformed to matrix form as follows:

$$\begin{aligned}
\bar{\Omega}_{ff} \dot{u}_f + \bar{\Omega}_{fl} \dot{u}_l^* &= \\
&\quad - k_q (\bar{\Omega}_{ff} q_f + \bar{\Omega}_{fl} q_l^*) - k_u (\bar{\Omega}_{ff} u_f + \bar{\Omega}_{fl} u_l^*) \\
&\quad + (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) + (\alpha \otimes I_n) \varepsilon(X) \\
&= -k_q \bar{\Omega}_{ff} (q_f + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} q_l^*) \\
&\quad - k_u \bar{\Omega}_{ff} (u_f + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} u_l^*) \\
&\quad + (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) + (\alpha \otimes I_n) \delta(X)
\end{aligned} \tag{8.12}$$

Considering (8.3) and (8.4), (8.12) becomes:

$$\begin{aligned}
\bar{\Omega}_{ff} \dot{u}_f + \bar{\Omega}_{fl} \dot{u}_l^* &= -k_q \bar{\Omega}_{ff} \varepsilon_{q_f} - k_u \bar{\Omega}_{ff} \varepsilon_{u_f} \\
&\quad + (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) + (\alpha \otimes I_n) \delta(X)
\end{aligned} \tag{8.13}$$

Equation (8.13) can be rewritten as:

$$\begin{aligned}\dot{u}_f = & -k_q \varepsilon_{q_f} - k_u \varepsilon_{u_f} - \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} \dot{u}_l^* \\ & + \bar{\Omega}_{ff}^{-1} (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) + \bar{\Omega}_{ff}^{-1} (\alpha \otimes I_n) \delta(X)\end{aligned}\quad (8.14)$$

From (8.4), one gets:

$$\begin{aligned}\dot{\varepsilon}_{u_f} = & \dot{u}_f + \bar{\Omega}_{ff}^{-1} \bar{\Omega}_{fl} \dot{u}_l^* \\ = & -k_q \varepsilon_{q_f} - k_u \varepsilon_{u_f} + \bar{\Omega}_{ff}^{-1} (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) \\ & + \bar{\Omega}_{ff}^{-1} (\alpha \otimes I_n) \delta(X)\end{aligned}\quad (8.15)$$

Theorem 8.1 *Suppose the MAS (8.1) satisfies the conditions of affine localizability detailed in [21]. If the distributed control protocol is designed as (8.6), with disturbance estimator (8.7) and neural network update law (8.8), then the closed-loop system is uniformly-ultimately bounded and the followers can successfully track the maneuvers determined by the leaders.*

Proof.

A positive definite function is selected as follows:

$$L = \frac{1}{2} \varepsilon_{u_f}^T \bar{\Omega}_{ff} \varepsilon_{u_f} + \frac{1}{2} \varepsilon_{q_f}^T \bar{\Omega}_{ff} \varepsilon_{q_f} + \frac{1}{2\sigma} \widetilde{W}^T \widetilde{W} \quad (8.16)$$

The derivative of L with respect to time gives: |

$$\begin{aligned}
\dot{L} &= \varepsilon_{u_f}^T \bar{\Omega}_{ff} \dot{\varepsilon}_{u_f} + \dot{\varepsilon}_{q_f}^T \bar{\Omega}_{ff} \varepsilon_{q_f} + \frac{1}{\sigma} \dot{\tilde{W}}^T \tilde{W} \\
&= \varepsilon_{u_f}^T \bar{\Omega}_{ff} \dot{\varepsilon}_{u_f} + \varepsilon_{u_f}^T \bar{\Omega}_{ff} \varepsilon_{q_f} - \frac{1}{\sigma} \dot{\tilde{W}}^T \tilde{W}
\end{aligned} \tag{8.17}$$

Evaluating (8.17) yields:

$$\begin{aligned}
\dot{L} &= \varepsilon_{u_f}^T \left(-\bar{\Omega}_{ff} \varepsilon_{q_f} - k_u \bar{\Omega}_{ff} \varepsilon_{u_f} + (\alpha \otimes I_n) \widetilde{W}^T \Phi(X) \right. \\
&\quad \left. + (\alpha \otimes I_n) \delta(X) \right) + \varepsilon_{u_f}^T \bar{\Omega}_{ff} \varepsilon_{q_f} - \varepsilon_{u_f}^T (\alpha \otimes I_n) \tilde{W}^T \Phi(X) \\
&= -k_u \varepsilon_{u_f}^T \bar{\Omega}_{ff} \varepsilon_{u_f} + \varepsilon_{u_f}^T (\alpha \otimes I_n) \delta(X)
\end{aligned} \tag{8.18}$$

By using the Frobenius norm, (8.18) becomes:

$$\begin{aligned}
\dot{L} &\leq -k_u \lambda_{\min}(\bar{\Omega}_{ff}) \|\varepsilon_{u_f}\|^2 + \|\varepsilon_{u_f}\| \eta \bar{\delta} \\
&\leq -\|\varepsilon_{u_f}\| \left(k_u \lambda_{\min}(\bar{\Omega}_{ff}) \|\varepsilon_{u_f}\| - \eta \bar{\delta} \right)
\end{aligned} \tag{8.19}$$

where $\eta = \|(\alpha \otimes I_n)\|$. It is obvious that $\dot{L} \leq 0$ provided that

$$k_u \lambda_{\min}(\bar{\Omega}_{ff}) \|\varepsilon_{u_f}\| - \eta \bar{\delta} > 0 \tag{8.20}$$

or

$$\|\varepsilon_{u_f}\| > \frac{\eta \bar{\delta}}{k_u \lambda_{\min}(\bar{\Omega}_{ff})} \equiv b_\varepsilon \tag{8.21}$$

Hence, $\dot{L} \leq 0$ is satisfied outside the compact set $\|\varepsilon_{u_f}\| \leq b_\varepsilon$. Therefore, similar to [199], the closed-loop system is uniformly ultimately bounded. Choosing large enough

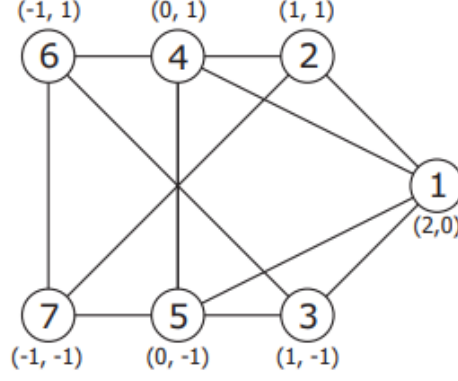


Figure 8.1: Undirected communication graph showing inter-agent communication.

k_u makes b_ε smaller and thus reduces the impact of the neural network approximation error in the closed-loop system.

8.4 Simulation results

This section presents a simulation example in order to show the effective performance of the neural network-based distributed controller. Here, seven quadrotors subjected to disturbances are considered as the agents. The nonlinear dynamic models of the quadrotors are described by the following equations [200]:

$$\ddot{x}_i = (C_{\phi_i} S_{\theta_i} C_{\psi_i} + S_{\phi_i} S_{\psi_i}) \frac{T_{1i}}{m_i} + d_{x_i} \quad (8.22)$$

$$\ddot{y}_i = (C_{\phi_i} S_{\theta_i} S_{\psi_i} - S_{\phi_i} C_{\psi_i}) \frac{T_{1i}}{m_i} + d_{y_i} \quad (8.23)$$

$$\ddot{z}_i = -g + (C_{\phi_i} C_{\theta_i}) \frac{T_{1i}}{m_i} + d_{z_i} \quad (8.24)$$

$$\ddot{\phi}_i = \dot{\theta}_i \dot{\psi}_i \frac{(I_{y_i} - I_{z_i})}{I_{x_i}} - J_p \dot{\theta}_i \varpi_i / I_{x_i} + \frac{T_{2i}}{I_{x_i}} + d_{\phi_i} \quad (8.25)$$

$$\ddot{\theta}_i = \dot{\phi}_i \dot{\psi}_i \frac{(I_{z_i} - I_{x_i})}{I_{y_i}} + J_p \dot{\phi}_i \varpi_i / I_{y_i} + \frac{T_{3i}}{I_{y_i}} + d_{\theta_i} \quad (8.26)$$

$$\ddot{\psi}_i = \dot{\phi}_i \dot{\theta}_i \frac{(I_{x_i} - I_{y_i})}{I_{z_i}} + \frac{T_{4i}}{I_{z_i}} + d_{\psi_i} \quad (8.27)$$

where x_i, y_i , and z describe the positions of the quadrotors with respect to the inertial reference frame, ϕ_i, θ_i , and ψ_i are the Euler angles representing the roll, pitch, and yaw of the quadrotor orientation, respectively, I_{x_i}, I_{y_i} , and I_{z_i} are the moments of inertia of the quadrotors, $\varpi_i = -\omega_{1i} + \omega_{2i} - \omega_{3i} + \omega_{4i}$ is the net speed of the rotors, $d_{x_i}, d_{y_i}, d_{z_i}, d_{\phi_i}, d_{\theta_i}$ and d_{ψ_i} are the adversarial disturbances, T_{1i}, T_{2i}, T_{3i} and T_{4i} are the control signals.

The dynamic equations describing the 2-D motion of the quadrotors can be extracted as follows [191], [190]:

$$\begin{cases} \ddot{x}_i = v_{x_i} + d_{x_i} \\ \ddot{y}_i = v_{y_i} + d_{y_i} \end{cases} \quad i \in \mathcal{V}_f \quad (8.28)$$

where

$$\begin{aligned} v_{x_i} &= (C_{\phi_i} S_{\theta_i} C_{\psi_i} + S_{\phi_i} S_{\psi_i}) \frac{T_{1i}}{m_i} \\ v_{y_i} &= (C_{\phi_i} S_{\theta_i} S_{\psi_i} - S_{\phi_i} C_{\psi_i}) \frac{T_{1i}}{m_i} \end{aligned}$$

Let $q_i = [x_i \ y_i]^T$, $u_i = [x_i \ y_i]^T$, $v_i = [v_{x_i} \ v_{y_i}]^T$ and $d_i = [d_{x_i} \ d_{y_i}]^T$, then

$$\begin{cases} \dot{q}_i = u_i \\ \dot{u}_i = v_i + d_i \end{cases} \quad i \in \mathcal{V}_f \quad (8.29)$$

where

$$d_i = \begin{bmatrix} 0.5x_i \cos(0.33 * t) \\ 0.81y_i \sin(0.91t) \end{bmatrix} \quad (8.30)$$

It should be noted that the positions, as well as the velocities of the leaders, are known in advance. The communication topology of the quadrotors together with their corresponding nominal positions is depicted in Fig. 8.1. The nominal positions of the quadrotors are $s_1 = [2 \ 0]$, $s_2 = [1 \ 1]$, $s_3 = [1 \ -1]$, $s_4 = [0 \ 1]$, $s_5 = [0 \ -1]$, $s_6 = [-1 \ 1]$ and $s_7 = [-1 \ -1]$ and $s = [s_1^T, s_2^T, \dots, s_7^T]^T$. The quadrotors labeled 1, 2, and 3 are the leaders while the rest are the followers. The stress matrix of the communication topology is calculated as [39]:

$$\bar{\Omega} = \begin{bmatrix} \bar{\Omega}_{ll} & \bar{\Omega}_{lf} \\ \bar{\Omega}_{fl} & \bar{\Omega}_{ff} \end{bmatrix} \quad (8.31)$$

where

$$\begin{aligned}
\bar{\Omega}_{ll} &= \begin{bmatrix} 0.2742 & -0.2741 & -0.2741 \\ -0.2741 & 0.6853 & 0 \\ -0.2741 & 0 & 0.6853 \end{bmatrix}, \\
\bar{\Omega}_{ff} &= \begin{bmatrix} 0.7538 & -0.0685 & -0.2741 & 0 \\ -0.0685 & 0.7538 & 0 & -0.2741 \\ -0.2741 & 0 & 0.2741 & -0.1370 \\ 0 & -0.2741 & -0.1370 & 0.2741 \end{bmatrix} \\
\bar{\Omega}_{fl} &= \begin{bmatrix} 0.1370 & -0.5482 & 0 \\ 0.1370 & 0 & -0.5482 \\ 0 & 0 & 0.1370 \\ 0 & 0.1370 & 0 \end{bmatrix} \\
\bar{\Omega}_{lf} &= \begin{bmatrix} 0.1370 & 0.1370 & 0 & 0 \\ -0.5482 & 0 & 0 & 0.1370 \\ 0 & -0.5482 & 0.1370 & 0 \end{bmatrix}
\end{aligned}$$

The gains of the proposed controller are set as $k_q = 10$, $k_v = 16$, $\sigma = 0.5$, The activation function is given by

$$\Pi_i(X_i) = \frac{1}{1 + e^{-X_i}}. \quad (8.32)$$

the input vector is set chosen as $X_i = [\varepsilon_{qf,i} \ \varepsilon_{uf,i}]^T$, the number of nodes of the hidden layer is set as 5, the initial condition of each neural network weight is set as 0.01.

The two-dimensional affine formation maneuvers of the seven quadrotors is pre-

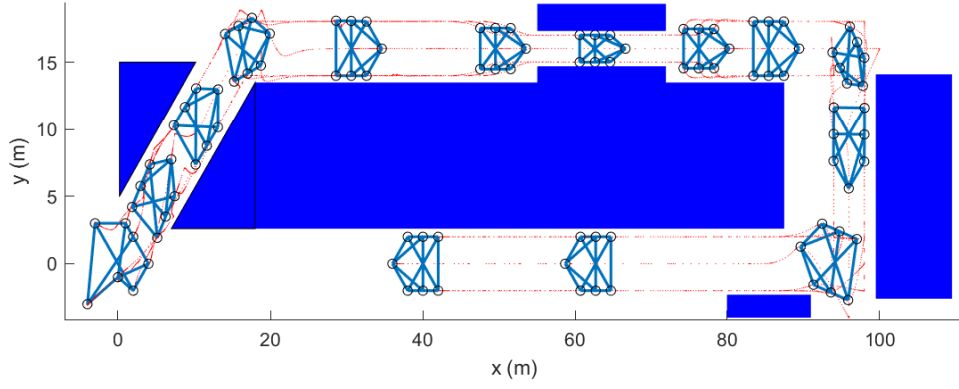


Figure 8.2: Two-dimensional affine formation maneuvering of the seven quadrotors.

sented in Fig. 8.2. It can be seen that the followers are tracking the trajectories of the leaders during translation, rotation, and scaling depending on the obstacle. Therefore, the proposed control method is able to suppress the effects of the adversarial disturbances and ensures the leader-follower affine formation maneuvers.

It is worth noting that the proposed control method in this paper is limited to second-order MASs in two-dimension but can be extended to three-dimension.

CHAPTER 9

CRITIC LEARNING-BASED ZERO-SUM DIFFERENTIAL GAME FOR AFFINE FORMATION MANEUVER OF MULTI-AGENT SYSTEMS WITH ACTUATOR ATTACKS

In this chapter, the affine formation maneuver (AFM) control problem using the zero-sum differential game is studied for a second-order multi-agent system (MAS) against actuator cyber-attacks. First and foremost, the distributed control signals and the cyber-injected attack signals are viewed as two competing teams. Then, the leader-follower tracking error system is established where the followers aim to track the AFMs of the leaders. The leader-follower distributed optimal control policies and the cyber-attack attenuation policies are derived from the Nash equilibrium solution of the Hamilton-Jacobi-Isaac (HJI) equation. The solution of the HJI equation is estimated with the aid of the critic neural network. The weights of the critic system are learned online based on the adaptive dynamic programming (ADP) scheme. The uniformly ultimately bounded stability of the closed-loop system under the formulated zero-sum differential game control strategy has been proven by the Lyapunov stability theorem. A simulation example illustrates that the presented control method can attenuate actuator cyber-attacks and maintain the AFMs of the agents.

9.1 Introduction

In practice, each agent in networked systems has limited fuel or energy, such as robot power, vehicle gasoline, and so on. In many engineering applications, reaching the target formation while using the least amount of energy is essential. As a result, optimal formation control is of great feasible and economic importance. In [65], an inverse optimality tool is used to calculate the optimal control policy to minimize the fuel consumption of the followers for cost-effective affine formation following. In [60], a linear quadratic regulator approach (LQR) scheme is developed to drive the agents to procure the target formation. In [201], a fully distributed optimal AFM control is designed for uncertain linear dynamic followers to track the stationary formation transformations of the leaders. The authors in [43] proposed a mixed integer semi-definite programming algorithm to minimize the communication cost and maximize the convergence speed of single-integrator leader-follower affine formation shape control. Regarding the low volume of reported results on optimal affine formation control, it is safe to say that the area is still open for more investigation.

The inter-agent interaction over communication networks exposed them to cyber-attacks from adversaries [202]. These attacks including replay attacks, injection attacks, and denial-of-service attacks may cause the failure of the AFM control systems and eventual collapse of the system [83], [203]. Therefore, protecting systems against cyber-attacks through effective control design is an interesting research topic and is garnering enormous attention. In some cases, the adversaries tend to sabotage the control systems by injecting false data into the actuators of the systems. Therefore,

securing the actuators against false data injection attacks should be of utmost importance [204]. So far, secure formation shape control of MASs against actuator attacks has been investigated [205], [206], [176], but the secure AFM control of MASs against actuator attacks received little attention [207]. Therefore, more attention should be given to protection against actuator attacks.

Though, the research results [205], [206], [176] simultaneously guaranteed formation control and attenuation of actuator attacks, the optimality of the system performance and cost-effective control were not considered. To address multi-objective optimization control problems, the zero-sum differential game is critical. In optimal control applications, there are normally two opposing players and the total reward of the game is zero in sum. The solution of the zero-sum differential game is derived from the Hamilton-Jacobi-Isaac (HJI) equation. Nevertheless, the analytical solution of the Hamilton-Jacobi-Isaac equation is difficult to obtain especially for complex networked systems. In order to tackle this difficulty, the solution of the HJI is approximated with the critic network. Adaptive dynamic programming (ADP) is introduced to estimate the weights of the critic network through policy iteration [208], value iteration [209], [210], or neural network estimation [211]. In [211], the zero-sum differential game theory is proposed for a microgrid system to mitigate disturbance and achieve optimal load frequency control. Gao et al. [212] proposed a robust adaptive optimal sliding mode control to guide a missile system with external disturbances. In [213] the authors studied the pursuers-evader control via the zero-sum differential game. In [214], the problem of optimal fault-tolerant control of discrete-time nonlinear systems is studied. The application of zero-sum differential game theory for secure control of

cyber-physical systems is studied in [215], [216], [217].

Nonetheless, in the framework of zero-sum differential game theory, there is very little literature on the security control of a single cyber-physical system subjected to false data injection attacks. Moreover, how to extend this to AFM control of multiple cyber-physical systems interacting over an undirected graph network is the motivation of this work.

Inspired by the foregoing discussion, this chapter introduces a zero-sum differential game for anti-attack optimal distributed AFM control of second-order leader-follower MASs facing actuator cyber-attacks. In this chapter, the distributed AFM control signals and the false data injection signals are treated as two opposing players. A single critic network is used to obtain the approximate optimal AFM control policy and the worst-case attack policy through the approximate solution of the HJI equation. Then, the ADP is employed to learn the weights of the critic network in real time. The proposed anti-attack control strategy can diminish the effects of the actuator attacks, and ensure the Leader-follower AFM of the compromised MAS with asymptotic stability. Finally, the contribution of this chapter is summarized in the following points.

1. To the best of our knowledge, this is the first time zero-sum differential game theory has been employed for optimal AFM control of MASs under false data injection attacks on the actuators.
2. In contrast to the optimal affine formation shape control in [65], [60], [201], [43] where the leaders are at rest, the leaders in this article are non-stationary can

achieve different maneuvers under the proposed optimal AFM control policy, which is more applicable in practice, especially in environments with obstacles.

3. Different from [207], [69], [43] where the controller gains are fixed, intelligent critic learning is used to learn the optimal control laws in real-time. This allows the control laws to adapt and adjust in response to various maneuvers and actuator attacks.
4. Contrary to [207] where two action networks and a critic network were used for accomplishing the control objectives, in this paper, a lone critic network is used to attain the control objectives. Thus, the proposed control approach is simpler and easy to apply to real-world multi-agent systems.
5. The critic network in [207] is constructed using a binary-based cost function which can induce a chattering effect as a result of switching between "0" and "1". On the contrary, this article uses a continuous quadratic cost function coupled with the HJI equation to actualize the optimal control policy.

This chapter is organized as follows: The problem statement and preliminaries are provided in Section 9.2. The secure control technique is designed in Section 9.3. The validity of the proposed control protocol is demonstrated in Section 9.4.

9.2 Problem formulation

9.2.1 Dynamics of the agents

A group of N mobile agents in two-dimensional space is considered in this study. Define two sets of integers $S_l = \{1, 2, \dots, N_l\}$ and $S_f = \{N_l + 1, N_l + 2, \dots, N\}$ as the leaders' subset and followers' subset, respectively.

The dynamics of the leaders are defined as follows:

$$\begin{aligned}\dot{q}_i &= v_i \\ \dot{v}_i &= u_i, \quad i \in \mathbb{S}_l\end{aligned}\tag{9.1}$$

where $q_i \in \mathbb{R}^n$, $v_i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}^n$ represent the position, velocity and control input of i th leader.

Expressing (9.1) in global form gives [15]:

$$\begin{aligned}\dot{q}_l &= v_l \\ \dot{v}_l &= u_l\end{aligned}\tag{9.2}$$

where $q_l = [q_1^T \ q_2^T \ \dots \ q_{N_l}^T]^T$, $v_l = [v_1^T \ v_2^T \ \dots \ v_{N_l}^T]^T$, $u_l = [u_1^T \ u_2^T \ \dots \ u_{N_l}^T]^T$.

The dynamics of the i th follower with compromised actuators is given by:

$$\begin{aligned}\dot{q}_i &= v_i \\ \dot{v}_i &= \check{u}_i, \quad i \in \mathbb{S}_f\end{aligned}\tag{9.3}$$

where $q_i \in \mathbb{R}^n$, $v_i \in \mathbb{R}^n$ and $\check{u}_i \in \mathbb{R}^n$ represent the position, velocity and compromised input of i th follower.

The comprised actuator as a result of the cyber-attack is modeled as follows [218]:

$$\check{u}_i = u_i + \Delta_i(t, q_i(t), v_i(t)) \quad (9.4)$$

where $u_i \in \mathbb{R}^n$ is the uncorrupted control input, and $\Delta_i(t, q_i(t), v_i(t)) \in \mathbb{R}^n$ denotes the injected attack.

The compromised follower system under the malicious actuator attack is given by:

$$\begin{aligned} \dot{q}_i &= v_i \\ \dot{v}_i &= u_i + \Delta_i(t, q_i(t), v_i(t)), \quad i \in \mathbb{S}_f \end{aligned} \quad (9.5)$$

The global form of (9.5) is thus:

$$\begin{aligned} \dot{q}_f &= v_f \\ \dot{v}_f &= u_f + \Delta_f \end{aligned} \quad (9.6)$$

where $q_f = [q_{N_l+1}^T \ q_{N_l+1}^T \cdots q_{N_l+N_f}^T]^T$, $v_f = [v_{N_l+1}^T \ v_{N_l+1}^T \cdots v_{N_l+N_f}^T]^T$, $u_f = [u_{N_l+1}^T \ u_{N_l+1}^T \cdots u_{N_l+N_f}^T]^T$, $\Delta_f = [\Delta_{N_l+1}^T \ \Delta_{N_l+2}^T \cdots \Delta_{N_l+N_f}^T]^T$

Assumption 9.1 [49], [47] *Let q_l^r be the required position of the leaders and u_l^r be the control protocol required to achieve the desired formation of the leaders. In this*

article, we assume that $q_l = q_l^r$ and $u_l = u_l^r$ for all t . Therefore, the control objective is now to realize $q_f = q_f^r$ as $t \rightarrow \infty$ in the presence of compromised actuators.

Remark 9.1 Due to the existence of cyber-attacks on the actuators of (9.3), the control protocol cannot be applied directly. The main aim of this article is to develop a secure distributed control capable of tolerating cyber-attacks, and driving the agents to the desired formation maneuver.

9.3 Control Design

This section introduces a zero-sum differential game through ADP to design a secure optimal control strategy for the AFM of the leaders (9.2) and followers (9.6). As mentioned in Assumption 9.1, the control of the leaders (9.2) will not be discussed. The control objective is to guarantee that the system (9.6) under contingencies can successfully track the AFM of the leaders. The AFM tracking errors of the followers are defined by:

$$\xi_i = \sum_{j=1}^N h_{ij}(q_i - q_j) \quad i \in \mathcal{V}_f \quad (9.7)$$

Expressing (9.7) into compact form gives:

$$\begin{aligned} \xi_q &= \bar{\Omega}_{ff} q_f + \bar{\Omega}_{fl} q_l = \bar{\Omega}_{ff} (q_f + (H_{ff}^{-1} H_{fl} \otimes I_n) q_l) \\ &= \bar{\Omega}_{ff} (q_f - q_f^r) \end{aligned} \quad (9.8)$$

where $\xi_q = [\xi_{N_l+1}^T \ \xi_{N_l+1}^T \ \dots \ \xi_{N_l+N_f}^T]^T$, and $q_f^r = -(H_{ff}^{-1} H_{fl} \otimes I_n) q_l$.

To derive the error dynamics, (9.8) is differentiated with respect to time, and using (9.6) yields:

$$\begin{cases} \dot{\xi}_q = \bar{\Omega}_{ff} (\dot{q}_f - \dot{q}_f^r) = \xi_v \\ \dot{\xi}_v = \bar{\Omega}_{ff} (u_f + \Delta_f - \ddot{q}_f^r) \end{cases} \quad (9.9)$$

Let $\delta_f = \Delta_f - \ddot{q}_f^r$, then, (9.9) can be rewritten as follows:

$$\dot{\xi} = A\xi + B\bar{\Omega}_{ff}u_f + C\bar{\Omega}_{ff}\delta_f \quad (9.10)$$

where $\xi = [\xi_q^T \ \xi_v^T]^T \in \mathbb{R}^{2N_f \times 2N_f}$,

$$A = \begin{pmatrix} 0_{N_f \times N_f} & I_{N_f \times N_f} \\ 0_{N_f \times N_f} & 0_{N_f \times N_f} \end{pmatrix}; \quad B = C = \begin{pmatrix} 0_{N_f \times N_f} \\ I_{N_f \times N_f} \end{pmatrix}$$

Define an infinite horizon cost function as follows:

$$J = \int_0^\infty \left(\xi^T Q \xi + u_f^T R u_f - \mu^2 \delta_f^T \delta_f \right) dt \quad (9.11)$$

where $Q = Q^T$ and $R = R^T$ are design matrices, $\mu > 0$ is the attenuation factor.

The zero-sum game issue aims to obtain a feedback control policy u_f such that the closed-loop system is asymptotically stable and its norm-2 is not greater than μ .

$$\int_0^\infty \left(\xi^T Q \xi + u_f^T R u_f \right) dt \leq \mu^2 \int_0^\infty \delta_f^T \delta_f dt \quad (9.12)$$

We want to find a saddle point solution (u_f^*, δ_f^*) , that satisfies the condition of

Nash equilibrium

$$J(\xi(0), u_f^*, \delta_f) \leq J(\xi(0), u_f^*, \delta_f^*) \leq J(\xi(0), u_f, \delta_f^*) \quad (9.13)$$

which is the typical zero-sum game. For the feedback admissible control pair $u_f^*(\xi(t))$ and $\delta_f(\xi(t))^*$, the optimal value function is given by:

$$V^*(\xi) = J(\xi(0), u_f^*, \delta_f^*) \quad (9.14)$$

When (9.14) is finite and continuously differentiable, the following relation is obtained:

$$\begin{aligned} 0 = & \xi^T Q \xi + u_f^T R u_f - \mu^2 \delta_f^T \delta_f \\ & + (\nabla V^*(\xi))^T (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f) \end{aligned} \quad (9.15)$$

with $V(0) = 0$. The Hamiltonian function is thus:

$$\begin{aligned} F(\xi, u_f, \delta_f, \nabla V^*(\xi)) = & \xi^T Q \xi + u_f^T R u_f - \mu^2 \delta_f^T \delta_f \\ & + (\nabla V^*(\xi))^T (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f) \end{aligned} \quad (9.16)$$

where ∇ is the gradient operator and $\nabla V^*(\xi) = \partial V^*(\xi) / \partial \xi$.

The optimal value function $V^*(\xi)$ ensures that the Hamilton-Jacobi-Isaac equation (HIJ) $\min_u \max_{\delta_f} F(\xi, u_f, \delta_f, \nabla V^*(\xi)) = 0$ holds.

The stationary conditions for deriving the optimal control policy and the worst-

case disturbance policy are expressed as:

$$\begin{cases} \frac{\partial F}{\partial u_f} = 0 \\ \frac{\partial F}{\partial \delta_f} = 0 \end{cases} \quad (9.17)$$

which implies that

$$u_f^*(\xi) = -\frac{1}{2}R^{-1}B^T\bar{\Omega}_{ff}\nabla V^*(\xi) \quad (9.18)$$

$$\delta_f^*(\xi) = \frac{1}{2\mu^2}C^T\bar{\Omega}_{ff}\nabla V^*(\xi) \quad (9.19)$$

Substituting (9.18) and (9.19) into HJI equation yields

$$\begin{aligned} 0 = & \xi^T Q \xi + \nabla V^{*T} A \xi - \frac{R^{-1}}{4} \nabla V^{*T} B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} \nabla V^* \\ & + \frac{1}{4} \mu^{-2} \nabla V^{*T} C C^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} \nabla V^* \end{aligned} \quad (9.20)$$

Remark 9.2 *The zero-sum game needs the solution of the HJI equation (9.20) to obtain the optimal control policy (9.18) and the worst-case disturbance policy (9.19). However, the HJI equation is nonlinear PDE in terms V^* , and its analytical solution is hard to compute. To tackle this problem, ADP and the learning ability of the critic neural network will be introduced in the subsequent section.*

9.3.1 Approximate solution of the HJI Equation Using Critic Learning Mechanism

This section presents an ADP to estimate the optimal solution of the HJI equation using the critic neural network. Then, the critic learning mechanism is employed to

estimate the weights of the critic network.

By utilizing the universal approximation property of the neural networks, the optimal value function can be approximated as follows:

$$V^*(\xi) = \psi_c^T(\xi)W_c^* + \sigma_c(\xi) \quad (9.21)$$

where $\psi_c(\xi) \in \mathbb{R}^m$ is the activation function, $W_c^* \in \mathbb{R}^m$ is the ideal neural network weight, m is the number of neurons in the hidden layer, and $\sigma_c(\xi)$ is the approximation error.

The critic neural network is used to estimate the weights of (9.21) as:

$$\widehat{V}(\xi) = \psi_c^T(\xi)\widehat{W}_c \quad (9.22)$$

where $\widehat{W}_c$ is the estimate of W_c^* .

The gradients of (9.21) and (9.22) are expressed as:

$$\nabla V^*(\xi) = \nabla \psi_c^T(\xi)W_c^* + \nabla \sigma_c(\xi) \quad (9.23)$$

$$\nabla \widehat{V}(\xi) = \nabla \psi_c^T(\xi)\widehat{W}_c \quad (9.24)$$

Incorporating (9.24) into (9.18) and (9.19) yields:

$$u^*(\xi) = -\frac{1}{2}R^{-1}B^T\bar{\Omega}_{ff}\left(\nabla \psi_c^T(\xi)W_c^* + \nabla \sigma_c(\xi)\right) \quad (9.25)$$

$$\delta_f^*(\xi) = \frac{1}{2\mu^2}C^T\bar{\Omega}_{ff}\left(\nabla \psi_c^T(\xi)W_c^* + \nabla \sigma_c(\xi)\right) \quad (9.26)$$

Including the critic neural estimate into (9.25) and (9.26) gives:

$$\widehat{u}_f(\xi) = -\frac{1}{2}R^{-1}B^T\bar{\Omega}_{ff}\nabla\psi_c^T(\xi)\widehat{W}_c \quad (9.27)$$

$$\widehat{\delta}_f(\xi) = \frac{1}{2\mu^2}C^T\bar{\Omega}_{ff}\nabla\psi_c^T(\xi)\widehat{W}_c \quad (9.28)$$

where $\widehat{u}$ and $\widehat{\delta}$ are the estimate of u_f , δ_f and $W_c^{(i)}$, respectively.

Applying the neural network approximation (9.21) to the HJI equation (9.16) leads to:

$$\begin{aligned} F(\xi, u_f, \delta_f, W_c^*) &= \xi^T Q \xi + u_f^T R u_f - \mu^2 \delta_f^T \delta_f \\ &+ (\nabla \psi_c^T(\xi) W_c^*) (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f) \triangleq e_{cH} \end{aligned} \quad (9.29)$$

where $e_{cH} = -(\nabla \sigma_c)^T (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f)$

The estimate of the HJI equation (9.29) is given by:

$$\begin{aligned} \widehat{F}(\xi, u_f, \delta_f, \widehat{W}_c) &= \xi^T Q \xi + u_f^T R u_f - \mu^2 \delta_f^T \delta_f \\ &+ (\nabla \psi_c^T(\xi) \widehat{W}_c) (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f) \triangleq e_c \end{aligned} \quad (9.30)$$

The estimation error of the HJI equation is derived as:

$$\begin{aligned} F(\xi, u_f, \delta_f, W_c^*) - \widehat{F}(\xi, \widehat{u}_f, \delta_f, \widehat{W}_c) &= (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\ &\times (A \xi + B \bar{\Omega}_{ff} u_f + C \bar{\Omega}_{ff} \delta_f) = e_{cH} - e_c \end{aligned} \quad (9.31)$$

where $\widetilde{W}_c = W_c^* - \widehat{W}_c$ is the critic weight estimation error.

Then,

$$e_c = (\nabla \psi_c^T(\xi) \widetilde{W}_c) \left(A\xi + B\bar{\Omega}_{ff}u_f + C\bar{\Omega}_{ff}\delta_f \right) + e_{cH} \quad (9.32)$$

To derive the critic-weight update law, we introduce the following objective function:

$$E_c = \frac{1}{2} e_c^T e_c \quad (9.33)$$

To train the critic neural network to learn the weight $\hat{W}_c$ that will minimize (9.33), the gradient descent algorithm is applied, which gives:

$$\begin{aligned} \dot{\hat{W}}_c &= -\gamma_c \frac{\theta}{(1 + \theta^T \theta)^2} \frac{\partial E_c}{\partial \hat{W}_c} \\ &= -\gamma_c \frac{\theta}{(1 + \theta^T \theta)^2} \left(\theta^T \hat{W}_c + \xi^T Q \xi + \hat{u}_f^T R \hat{u}_f - \mu^2 \hat{\delta}_f^T \hat{\delta}_f \right) \end{aligned} \quad (9.34)$$

where $\gamma_c > 0$ is the learning rate, the term $(1 + \theta^T \theta)^2$ is used for normalization,

$$\theta = (\nabla \psi_c(\xi)) \left(A\xi + B\bar{\Omega}_{ff}\hat{u}_f + C\bar{\Omega}_{ff}\hat{\delta}_f \right)$$

The dynamics of the critic weight estimation error is derived as follows:

$$\begin{aligned} \dot{\widetilde{W}}_c &= -\gamma_c \frac{\theta}{(1 + \theta^T \theta)^2} \frac{\partial E_c}{\partial \widetilde{W}_c} \\ &= -\gamma_c \theta_1 \theta_1^T \widetilde{W}_c + \gamma_c \frac{\theta_1}{\theta_2} e_{cH} \end{aligned} \quad (9.35)$$

where

$$\theta_1 = \frac{\theta}{(1 + \theta^T \theta)}; \quad \theta_2 = 1 + \theta^T \theta \geq 1$$

9.3.2 Stability analysis of the closed-loop system

This section presents Lyapunov stability theory to demonstrate the stability analysis of the system (9.10) under the zero-sum game-based secure affine formation maneuver control.

Theorem 9.1 *Consider the leaders' subsystem (9.2), followers' subsystem (9.6), and the followers' tracking error system (9.10). Then, using the estimated control policy (9.27), the estimated worst-case disturbance policy (9.28), critic weights' update law (9.34), and the approximate value function (9.22), it is guaranteed that the closed-loop tracking errors are uniformly ultimately bounded, and various affine formation maneuvers can be attained by the agents.*

Proof. Select a Lyapunov function of the form:

$$\begin{aligned} L &= V^*(\xi) + \frac{1}{2} \widetilde{W}_c^T \widetilde{W}_c \\ &\triangleq L_1 + L_2 \end{aligned} \tag{9.36}$$

where $L_1 \triangleq V^*(\xi)$, and $L_2 \triangleq \frac{1}{2} \widetilde{W}_c^T \widetilde{W}_c$ ■

Differentiating L_1 with respect to time leads to:

$$\begin{aligned} \dot{L}_1 &= \frac{\partial V^*(\xi)}{\partial \xi} \dot{\xi} \\ &= (\nabla V^*(\xi))^T \left(A\xi + B\bar{\Omega}_{ff}\widehat{u}_f + C\bar{\Omega}_{ff}\widehat{\delta}_f \right) \\ &= (\nabla V^*(\xi))^T A\xi + (\nabla V^*(\xi))^T B\bar{\Omega}_{ff}\widehat{u}_f + (\nabla V^*(\xi))^T C\bar{\Omega}_{ff}\widehat{\delta}_f \end{aligned} \tag{9.37}$$

Note that (9.18), (9.19) and (9.20) yield:

$$(\nabla V^*(\xi))^T B \bar{\Omega}_{ff} = -2Ru^* \quad (9.38)$$

$$(\nabla V^*(\xi))^T C \bar{\Omega}_{ff} = 2\mu^2 \delta_f^* \quad (9.39)$$

$$\begin{aligned} (\nabla V^*(\xi))^T A \xi &= -\xi^T Q \xi - u_f^{*T} R u_f^* + \mu^2 \delta_f^{*T} \delta_f^* \\ &\quad - (\nabla V^*(\xi))^T B \bar{\Omega}_{ff} u_f^* - (\nabla V^*(\xi))^T C \bar{\Omega}_{ff} \delta_f^* \\ &= -\xi^T Q \xi - u_f^{*T} R u_f^* + \mu^2 \delta_f^{*T} \delta_f^* \\ &\quad + \frac{1}{2} R^{-1} (\nabla V^*(\xi))^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla V^*(\xi)) \\ &\quad - \frac{1}{2\mu^2} (\nabla V^*(\xi))^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla V^*(\xi)) \end{aligned} \quad (9.40)$$

Inserting (9.38), (9.39) and (9.40) into (9.37), one has:

$$\begin{aligned} \dot{L}_1 &= -\xi^T Q \xi - u_f^{*T} R u_f^* + \mu^2 \delta_f^{*T} \delta_f^* - 2u_f^{*T} R \hat{u}_f + 2\mu^2 \delta_f^{*T} \hat{\delta} \\ &\quad + \frac{1}{2} R^{-1} (\nabla V^*(\xi))^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla V^*(\xi)) \\ &\quad - \frac{1}{2\mu^2} (\nabla V^*(\xi))^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla V^*(\xi)) \end{aligned} \quad (9.41)$$

Utilizing (9.27) and (9.28), we can derive the following relations:

$$\begin{aligned} \hat{u}(\xi) &= -\frac{1}{2} R^{-1} B^T \bar{\Omega}_{ff} \nabla \psi_c^T(\xi) \widehat{W}_c + u^* - u^* \\ &= u^* + \frac{1}{2} R^{-1} B^T \bar{\Omega}_{ff} \nabla \psi_c^T(\xi) \widetilde{W}_c \\ &\quad + \frac{1}{2} R^{-1} B^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi)) \end{aligned} \quad (9.42)$$

$$\begin{aligned}
\widehat{\delta}(\xi) &= \frac{1}{2\mu^2} C^T \bar{\Omega}_{ff} \nabla \psi_c^T(\xi) \widehat{W}_c + \delta^* - \delta^* \\
&= \delta^* - \frac{1}{2\mu^2} C^T \bar{\Omega}_{ff} \nabla \psi_c^T(\xi) \widetilde{W}_c \\
&\quad - \frac{1}{2\mu^2} C^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi))
\end{aligned} \tag{9.43}$$

where $\widetilde{W}_c = W_c - \widehat{W}_c$

Using (9.42) and (9.43), one gets:

$$\begin{aligned}
u_f^{*T} R \widehat{u}_f &= u_f^{*T} R u_f^* \\
&\quad + \frac{1}{4} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\
&\quad + \frac{1}{4} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi))
\end{aligned} \tag{9.44}$$

$$\begin{aligned}
\delta_f^{*T} \widehat{\delta} &= \delta_f^{*T} \delta^* \\
&\quad - \frac{1}{4\mu^4} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\
&\quad - \frac{1}{4\mu^4} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi))
\end{aligned} \tag{9.45}$$

From (9.44) and (9.45), it follows that:

$$\begin{aligned}
\dot{L}_1 = & -\xi^T Q \xi - 3u_f^{*T} R u_f^* + 3\mu^2 \delta_f^{*T} \delta^* \\
& - \frac{1}{2} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\
& - \frac{1}{2\mu^2} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c^*) \\
& + \frac{1}{2} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) W_c^*) \\
& - \frac{1}{2\mu^2} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) W_c^*) \\
& + \frac{1}{4} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi)) \\
& - \frac{1}{4\mu^4} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi))
\end{aligned} \tag{9.46}$$

Further evaluating (9.46) leads to:

$$\begin{aligned}
\dot{L}_1 = & -\xi^T Q \xi \\
& - \frac{1}{2} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\
& - \frac{1}{2\mu^2} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) \widetilde{W}_c) \\
& - \frac{1}{4} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) W_c^*) \\
& - \frac{5}{4\mu^2} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \psi_c^T(\xi) W_c^*) \\
& + \frac{1}{4} R^{-1} (\nabla \psi_c^T(\xi) W_c^*)^T B B^T \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi)) \\
& - \frac{1}{4\mu^4} (\nabla \psi_c^T(\xi) W_c^*)^T C^T C \bar{\Omega}_{ff}^T \bar{\Omega}_{ff} (\nabla \sigma_c(\xi))
\end{aligned} \tag{9.47}$$

By utilizing the norm-2, (9.47) reduces to:

$$\begin{aligned}
\dot{L}_1 &\leq -\lambda_{\min}(Q) \|\xi\|^2 - \kappa_1 \|W_c^*\| \|\widetilde{W}_c\| - \kappa_2 \|W_c^*\| \|\widetilde{W}_c\| \\
&\quad - \kappa_3 \|W_c^*\|^2 - \kappa_4 \|W_c^*\|^2 + \frac{\kappa_3}{\zeta_{\psi_c}} \|W_c^*\| \zeta_{\sigma_c} - \frac{\kappa_4}{5\mu^2 \zeta_{\psi_c}} \|W_c^*\| \zeta_{\sigma_c} \\
&\leq -\lambda_{\min}(Q) \|\xi\|^2 - (\kappa_1 + \kappa_2) \|W_c^*\| \|\widetilde{W}_c\| \\
&\quad - (\kappa_3 + \kappa_4) \|W_c^*\|^2 + \left(\frac{\kappa_3}{\zeta_{\psi_c}} - \frac{\kappa_4}{5\mu^2 \zeta_{\psi_c}} \right) \zeta_{\sigma_c} \|W_c^*\|
\end{aligned} \tag{9.48}$$

where $\kappa_1 = \frac{1}{2} \lambda_{\min}(R^{-1}) \lambda_{\min}(\bar{\Omega}_{ff}) \zeta_B^2 \zeta_{\psi_c}^2$, $\kappa_2 = \frac{1}{2\mu^2} \lambda_{\min}(\bar{\Omega}_{ff}) \zeta_C^2 \zeta_{\psi_c}^2$, $\kappa_3 = \frac{1}{2} \kappa_1$, $\kappa_4 = \frac{5}{2} \kappa_2$, $\|B\|^2 \leq \zeta_B^2$, $\|C\|^2 \leq \zeta_C^2$, and $\|\nabla \psi_c^T(\xi)\|^2 \leq \zeta_{\psi_c}^2$, and $\|\nabla \sigma_c(\xi)\| \leq \zeta_{\sigma_c}$, with ζ_B , ζ_C , ζ_{ψ_c} , and ζ_{σ_c} being positive constants.

Noting that $\kappa_1 = 2\kappa_3$ and $\kappa_2 = 2/5\kappa_4$, and according to Young's inequality, we have:

$$\begin{aligned}
-(\kappa_1 + \kappa_2) \|W_c^*\| \|\widetilde{W}_c\| &= -\left(2\kappa_3 + \frac{2}{5}\kappa_4\right) \|W_c^*\| \|\widetilde{W}_c\| \\
&\leq \frac{\|\widetilde{W}_c\|^2}{2} + \left(2\kappa_3 + \frac{2}{5}\kappa_4\right)^2 \frac{\|W_c^*\|^2}{2}
\end{aligned} \tag{9.49}$$

By inserting the inequality in (9.49) into (9.48), we get:

$$\begin{aligned}
\dot{L}_1 &\leq -\lambda_{\min}(Q) \|\xi\|^2 + \frac{\|\widetilde{W}_c\|^2}{2} + \left(2\kappa_3 + \frac{2}{5}\kappa_4\right)^2 \frac{\|W_c^*\|^2}{2} \\
&\quad - (\kappa_3 + \kappa_4) \|W_c^*\|^2 + \left(\frac{\kappa_3}{\zeta_{\psi_c}} - \frac{\kappa_4}{5\mu^2 \zeta_{\psi_c}} \right) \zeta_{\sigma_c} \|W_c^*\|
\end{aligned} \tag{9.50}$$

Differentiating L_2 with respect to time gives:

$$\begin{aligned}
\dot{L}_2 &= \widetilde{W}_c^T \dot{\widetilde{W}}_c \\
&= -\gamma_c \widetilde{W}_c^T \theta_1 \theta_1^T \widetilde{W}_c + \gamma_c \frac{\widetilde{W}_c^T \theta_1}{\theta_2} e_{cH} \\
&\leq -\left(\gamma_c - \frac{1}{2}\right) \zeta_{\theta_1}^2 \|\widetilde{W}_c\|^2 + \frac{1}{2} \gamma_c^2 \zeta_{\theta_2 e_{cH}}^2
\end{aligned} \tag{9.51}$$

where $\|\theta\|^2 \leq \zeta_{\theta_1}^2$, and $\|e_{cH}/\theta_2\|^2 \leq \zeta_{\theta_2 e_{cH}}^2$.

The time-derivative of L is thus:

$$\begin{aligned}
\dot{L} &= \dot{L}_1 + \dot{L}_2 \\
&\leq -\lambda_{\min}(Q) \|\xi\|^2 - \left(\gamma_c - \frac{1}{2} - \frac{1}{2\zeta_{\theta_1}^2}\right) \zeta_{\theta_1}^2 \|\widetilde{W}_c\|^2 + O
\end{aligned} \tag{9.52}$$

where

$$\begin{aligned}
O &= \left(2\kappa_3 + \frac{2}{5}\kappa_4\right)^2 \frac{\|W_c^*\|^2}{2} + \frac{1}{2} \gamma_c^2 \zeta_{\theta_2 e_{cH}}^2 \\
&\quad - (\kappa_3 + \kappa_4) \|W_c^*\|^2 + \left(\frac{\kappa_3}{\zeta_{\psi_c}} - \frac{\kappa_4}{5\mu^2 \zeta_{\psi_c}}\right) \zeta_{\sigma_c} \|W_c^*\|
\end{aligned}$$

Given that the following inequalities hold:

$$O > 0; \quad \lambda_{\min}(Q) > 0 \quad ; \quad \left(\gamma_c - \frac{1}{2} - \frac{1}{2\zeta_{\theta_1}^2}\right) > 0 \tag{9.53}$$

and

$$\|\xi\| > \sqrt{\frac{O}{\lambda_{\min}(Q)}} \tag{9.54}$$

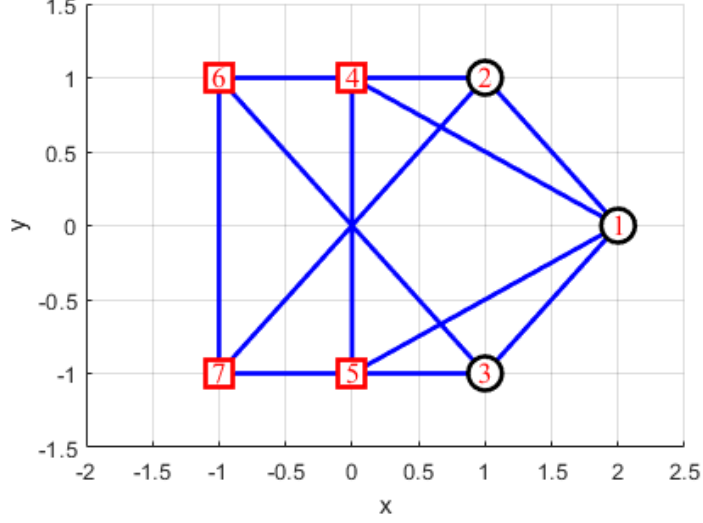


Figure 9.1: Target formation depicting the communication pattern and positions of the seven agents.

or

$$\|\widetilde{W}_c\| > \sqrt{O/\left(\gamma_c - \frac{1}{2} - \frac{1}{2\zeta_{\theta_1}^2}\right)} \quad (9.55)$$

then, $\dot{L} < 0$. Thus, the closed-loop signals including ξ and $\widetilde{W}_c$ are uniformly ultimately bounded.

9.4 Simulation results

This section presents numerical examples to verify the proposed secure distributed control laws.

The constant nominal formation $(\mathbb{G}, \varsigma)$ consisting of seven agents is depicted in Fig. 9.1. From this figure, the subset of the leaders and the followers are $S_l = \{1, 2, 3\}$ and $S_f = \{4, 5, 6, 7\}$, respectively. The nominal position of each agent is given by $\varsigma_1 = [2 \ 0]^T$, $\varsigma_2 = [1 \ 1]^T$, $\varsigma_3 = [1 \ -1]^T$, $\varsigma_4 = [0 \ 1]^T$, $\varsigma_5 = [0 \ -1]^T$, $\varsigma_6 = [-1 \ 1]^T$, and

$\varsigma_7 = [-1 \ -1]^T$, with $\varsigma = [\varsigma_1^T, \varsigma_2^T, \varsigma_3^T, \varsigma_4^T, \varsigma_5^T, \varsigma_6^T, \varsigma_7^T]^T$. The equilibrium stress matrix $\bar{\Omega}$

is computed based on the linear inequality matrix detailed in [21], Proposition 2, as:

$$\bar{\Omega} = \begin{bmatrix} \bar{\Omega}_{ll} & \bar{\Omega}_{lf} \\ \bar{\Omega}_{fl} & \bar{\Omega}_{ff} \end{bmatrix} \quad (9.56)$$

where

$$\bar{\Omega}_{ll} = \begin{bmatrix} 0.2742 & -0.2741 & -0.2741 \\ -0.2741 & 0.6853 & 0 \\ -0.2741 & 0 & 0.6853 \end{bmatrix},$$

$$\bar{\Omega}_{lf} = \begin{bmatrix} 0.1370 & 0.1370 & 0 & 0 \\ -0.5482 & 0 & 0 & 0.1370 \\ 0 & -0.5482 & 0.1370 & 0 \end{bmatrix}$$

$$\bar{\Omega}_{fl} = \begin{bmatrix} 0.1370 & -0.5482 & 0 \\ 0.1370 & 0 & -0.5482 \\ 0 & 0 & 0.1370 \\ 0 & 0.1370 & 0 \end{bmatrix}$$

$$\bar{\Omega}_{ff} = \begin{bmatrix} 0.7538 & -0.0685 & -0.2741 & 0 \\ -0.0685 & 0.7538 & 0 & -0.2741 \\ -0.2741 & 0 & 0.2741 & -0.1370 \\ 0 & -0.2741 & -0.1370 & 0.2741 \end{bmatrix}$$

Let the actuator attack signals be $\delta_i(t, p_i, v_i) = [2\cos(q_i^2 t), 2\sin(v_i^2 t)]^T$. It is assumed that the attackers targeted the actuators of agent-4 and agent-6 for a period of $20 \text{ s} < t \leq 30 \text{ s}$. We set $Q = 10I_{8 \times 8}$, $R = 20I_{4 \times 4}$, $\mu = 5$, the basis functions and the weight vectors are chosen as:

$$\psi_{ci} = [\xi_i^T \xi_i, \xi_i^T \dot{\xi}_i, \dot{\xi}_i^T \dot{\xi}_i] \in \mathbb{R}^{1 \times 6} \quad (9.57)$$

$$W_{ci}^* = [w_{1i}^T, w_{2i}^T, w_{3i}^T]^T \in \mathbb{R}^{6 \times 1} \quad (9.58)$$

with

$$\nabla \psi_{ci} = \begin{bmatrix} \xi_i^T & \dot{\xi}_i^T & 0 \\ 0 & \xi_i^T & \dot{\xi}_i^T \end{bmatrix} \in \mathbb{R}^{2 \times 6} \quad (9.59)$$

Then the critic learning-based optimal control policy and the worst-case cyber-injection attack policy associated with each follower are given by:

$$\hat{u}_i = -\frac{1}{2}H_{ij}R_i^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} \xi_{pi}^T & 0 \\ \xi_{vi}^T & \xi_{pi}^T \\ 0 & \xi_{vi}^T \end{bmatrix}^T \hat{W}_{ci}, \quad i \in \mathbb{S}_f \quad (9.60)$$

$$\hat{\delta}_i = \frac{1}{2\mu^2}H_{ij}R_i^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} \xi_{pi}^T & 0 \\ \xi_{vi}^T & \xi_{pi}^T \\ 0 & \xi_{vi}^T \end{bmatrix}^T \hat{W}_{ci}, \quad i \in \mathbb{S}_f \quad (9.61)$$

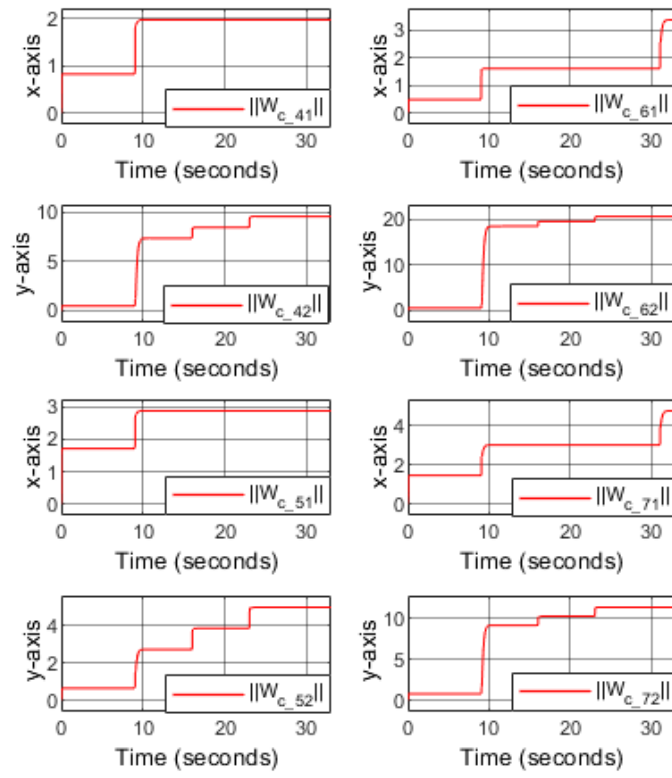


Figure 9.2: Evolution of the norm-2 of the Weights of the critic networks of the agents.

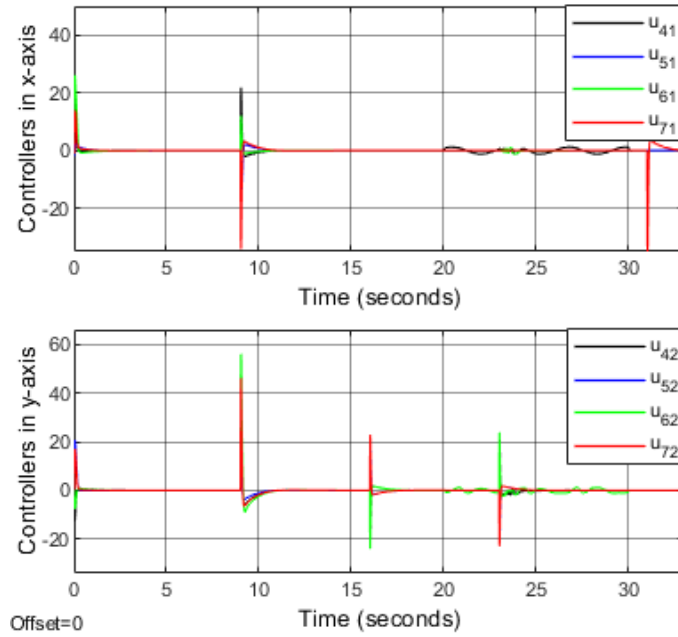


Figure 9.3: The attack-tolerant optimal control policies.

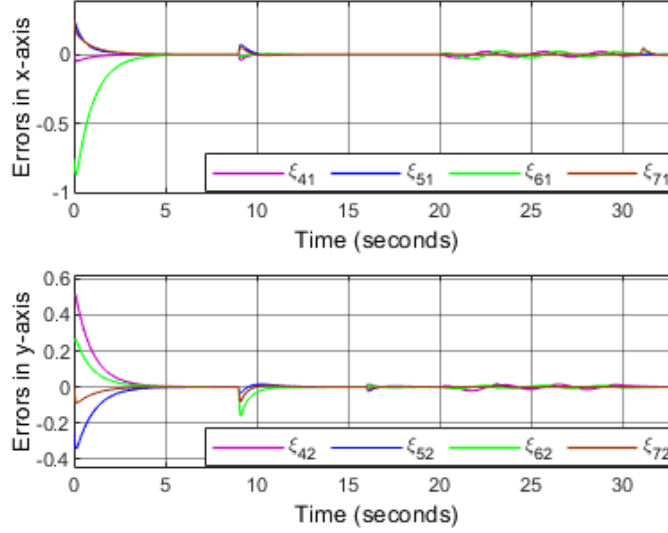


Figure 9.4: The tracking errors of the followers.

Figure (9.2) depicts the evolution of the weights of the critic networks of the followers. The critic weights are learned online according to the type of formation maneuver. The distributed optimal control policies that guarantee the AFM of the agents are shown in Fig. 9.3. The control laws resulted in the bounded tracking errors presented in Fig. 9.4. From Figs. 9.3 and 9.4, it can be seen that the actuator attacks on agent-4 and agent-6 resulted in some fluctuations which are very close to zero for $20\text{ s} < t \leq 30\text{ s}$. The AFMs of the followers and the leaders are shown in Fig. 9.5. This figure indicates that the proposed control method successfully attenuates the harmful effects of the actuator cyber-attacks on the leader-follower AFMs.

To highlight the merit of the proposed secure control method, the control approach developed in [68] is applied to the leaders' subsystem (9.2) and followers' subsystem (9.6). The simulation result is shown in Fig. 9.6. It can be observed that for the attack period $20\text{ s} < t \leq 30\text{ s}$, the formation of the agents in [68] is distorted. This is because the control method in [68] cannot counter the actuator cyber-attacks.

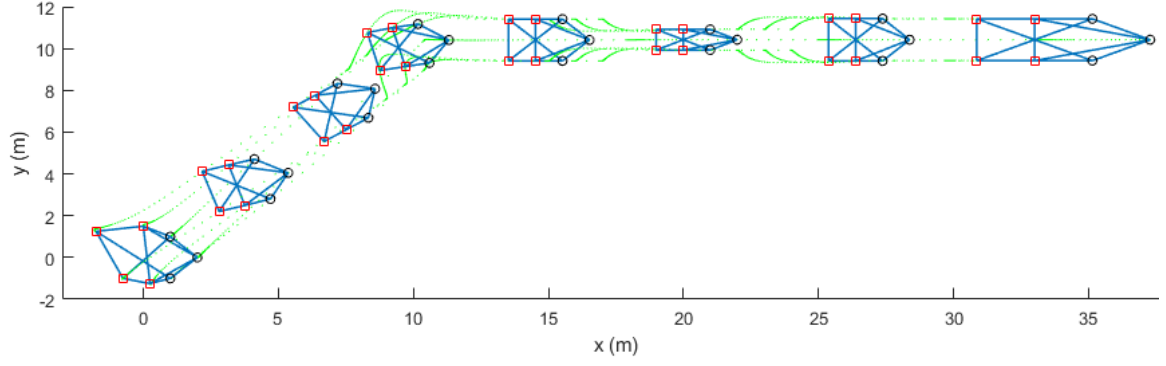


Figure 9.5: The affine transformations of the leader-follower nominal formation of the seven agents.

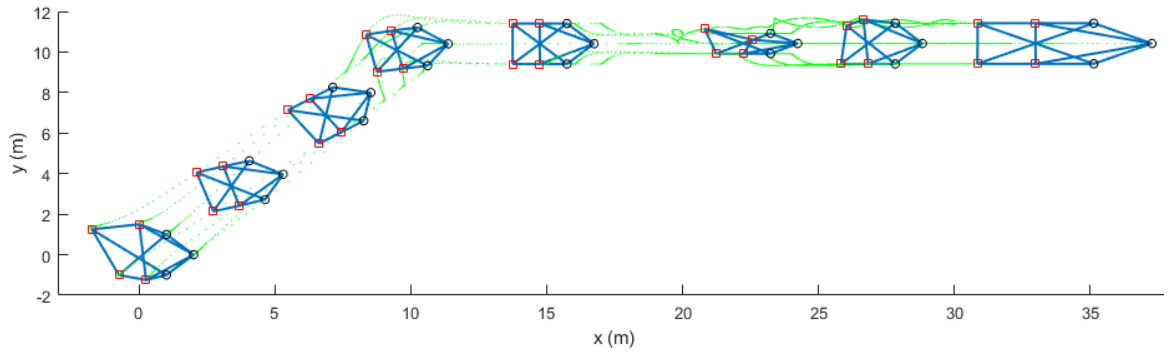


Figure 9.6: The affine formation maneuvers of the seven agents under the control approach in (Zhi et al. 2021).

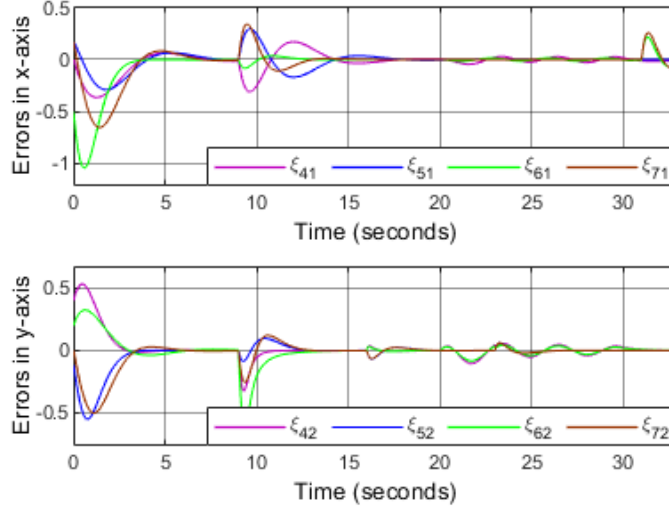


Figure 9.7: The tracking errors of the followers under the AFM control strategy proposed in (El-Ferik et al 2023).

To further demonstrate the advantage of the proposed secure AFM control, its tracking accuracy is compared with the recently published attack-tolerant AFM control discussed in [207]. This controller is capable of giving the same result as in Fig. 9.5. However, as shown in Fig. 9.7, the magnitudes of the tracking errors are bigger than in Fig. 9.4. The reason for this is that in [207], the authors used many approximations, including two action networks, a critic network, and the Nussbaum function. On the other hand, in this work, only the critic network is used. Therefore, the errors are smaller as depicted in Fig. 9.4. To highlight this fact, the mean-squared tracking errors of the proposed control technique and those in [207] are presented in Table 9.1

Table 9.1: Comparison of the mean-squared tracking errors of the proposed control technique and that in (El-Ferik et al 2023).

AFM tracking errors	Proposed control	Control in (El-Ferik et al 2023)
ξ_{41}	0.0007	0.0550
ξ_{51}	0.0039	0.0509
ξ_{61}	0.0014	0.0913
ξ_{71}	0.0036	0.0770
ξ_{42}	0.0083	0.0806
ξ_{52}	0.0052	0.0622
ξ_{62}	0.0047	0.0801
ξ_{72}	0.0014	0.0611

CHAPTER 10

CONCLUSIONS AND FUTURE WORK

This chapter provides a summary of the problems addressed in this thesis as well as the possible extension of the thesis in the future.

10.1 Conclusions

This thesis studied a series of control problems of secure affine formation maneuvers of multi-agent cyber-physical systems under cyber-attacks using learning mechanisms based on artificial intelligence. In practical cyber-physical systems, the exact system models are not available. Moreover, the magnitudes of external disturbances and cyber-attack signals that deteriorate the performance of cyber-physical systems are difficult to know in advance. With the aid of the learning mechanisms, the magnitudes of external disturbances and cyber-attack signals can be estimated and compensated through control policies.

In Chapter 4, the anti-attack leader-follower affine formation maneuver control of second-order multi-agent systems with actuator cyber-attacks has been studied. A reinforcement learning control algorithm based on actor-critic structure was used to ensure secure collective maneuvers of the agents. The Nussbaum functions are used to estimate the sign of the control laws since attacks can lead to a loss of control direction. A simulation example was given to demonstrate the effectiveness of the proposed control scheme.

In Chapter 5, to achieve attack mitigation and leader-follower tracking in finite time, the technique in Chapter 4 is extended. Here, a finite-time convergence theory and distributed fast terminal sliding mode control are used for affine formation maneuver control of nonlinear multi-agent systems subjected to sensor deception attacks and backlash-like hysteresis. The agents' nonlinearities, deception attacks, and backlash-like hysteresis are addressed via actor-critic reinforcement learning. A numerical simulation is provided to highlight the feasibility of the suggested control approach.

In Chapter 6, a prescribed performance function was used to predefine the settling time, and transient and steady states of the agents subjected to sensor deception attacks and actuator faults. The secure affine formation maneuvers of the agents were realized using an optimized secure backstepping control with prescribed performance. Then, a reinforcement learning approach based on the identifier-actor-critic architecture is employed to address the issue of the unknown dynamics, deception attacks,

and faults. The simulation results are presented to validate the control scheme.

In Chapter 7, the resilient and safe leader-follower affine formation maneuver of a group of quadrotors under external disturbances, sensor deception attacks, and actuator injection attacks is presented. Considering the physical limitations of quadrotors, their dynamic states were constrained within a safety boundary, and a barrier Lyapunov function was used to prevent the quadrotors from breaching the constraints. The unknown adversaries confronting the quadrotors were neutralized using a reinforcement learning scheme and a new distributed sliding mode control. Simulation results demonstrate the efficacy of the designed resilient control protocol.

In Chapter 8, the resilient affine formation maneuver control of nonlinear multi-agent systems with adversarial disturbances was provided. Here, an artificial neural network is employed as an intelligent framework to detect, estimate, and mitigate harmful adversarial disturbances. Then, a nonlinear control technique based on the variable structure was utilized to achieve a resilient leader-follower collective formation maneuvering despite the disturbances. The proposed control algorithm was applied to a system of multiple quadrotor UAVs. Simulation results have been provided to highlight the advantage of the resilient control approach.

In Chapter 9, the affine formation maneuver control problem of a second-order multi-agent system under actuator attacks was investigated. the distrusted optimal control policy and the actuator attack signals are treated as two opposing teams.

Then, a zero-sum differential game algorithm which was solved through the Nash equilibrium solution of the HJI was employed to derive the distributed optimal control policies and the cyber-attack attenuation policies. The solution of the HJI equation was estimated with the aid of the critic neural network. The weights of the critic system are learned online based on the ADP scheme. Numerical examples were provided to validate the performance of the proposed control method.

10.2 Future Works

Affine formation maneuver control is still at an early stage compared to order formation control methods (such as bearing-based, distance-based, and position-based). Considering that the wireless communication among the agents to maintain the target affine formation maneuvers exposed them to cyber-attacks, this thesis proposed resilient learning-based control techniques to secure the agents. There are still problems to investigate in the future such as:

- In the modeling of practical multi-agent cyber-physical systems, the unmodeled parts are usually neglected. However, this may deteriorate the control performance and subsequently cause the failure of the system. In order to apply the affine formation maneuver control to practical systems, the unmodeled dynamics should be taken into account. Simultaneous consideration of unmodeled dynamics and cyber-attacks in designing affine formation maneuver control for multi-agent systems is significant.
- It is always desired to design coordinated control laws with limited communica-

tion resources. Even though event-triggered control has been reported for affine formation maneuver control to minimize communication frequency and energy [64], [72], it has not been used together with learning algorithms to counter cyber-attacks during affine formation maneuvers.

- In this thesis and existing affine formation maneuver control in the literature, it is assumed that the maneuvers are known in advance (i.e the path of the agents and the positions of obstacles along the path are known). Nonetheless, in practice, the obstacles can be dynamic, and unexpected obstructions might occur along the path. As a result, the agents tracking the predefined maneuvers cannot autonomously avoid the dynamic obstacles and unexpected obstructions. Therefore, a potential field function should be combined with affine formation maneuver control theory as a novel approach so that the agents can autonomously avoid obstacles. This will eliminate the common assumption that the maneuvers are known in advance.
- In this thesis and existing literature on affine formation maneuver control, for any n -dimensional space, the number of leaders must be at least $n+1$ to satisfy the condition of affine localizability. This condition is too strict and has ruled out the application of affine formation maneuver control in situations where the number of leaders is less than $n+1$. Thus, how to design affine formation maneuver control with one leader or less than $n+1$ leaders is an interesting research direction.

REFERENCES

- [1] R. Owoputi and S. Ray, “Security of multi-agent cyber-physical systems: A survey,” *IEEE Access*, vol. 10, pp. 121 465–121 479, 2022.
- [2] A. Amato, A. Quarto, and V. Di Lecce, “An application of cyber-physical system and multi-agent technology to demand-side management systems,” *Pattern recognition letters*, vol. 141, pp. 23–31, 2021.
- [3] T. Sanislav, S. Zeadally, G. Mois, and H. Fouchal, “Multi-agent architecture for reliable cyber-physical systems (cps),” in *2017 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2017, pp. 170–175.
- [4] D. Ding, Q.-L. Han, Z. Wang, and X. Ge, “A survey on model-based distributed control and filtering for industrial cyber-physical systems,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2483–2499, 2019.
- [5] P. Pop, M. L. Raagaard, S. S. Craciunas, and W. Steiner, “Design optimisation of cyber-physical distributed systems using ieee time-sensitive networks,” *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 86–94, 2016.

- [6] Q. Jiao, H. Zhang, S. Xu, F. L. Lewis, and L. Xie, “Bipartite tracking of homogeneous and heterogeneous linear multi-agent systems,” *International Journal of Control*, vol. 92, no. 12, pp. 2963–2972, 2019.
- [7] L. Mo, Y. Niu, and T. Pan, “Consensus of heterogeneous multi-agent systems with switching jointly-connected interconnection,” *Physica A: Statistical Mechanics and its Applications*, vol. 427, pp. 132–140, 2015.
- [8] R. Wang, X. Dong, Q. Li, and Z. Ren, “Distributed time-varying formation control for multiagent systems with directed topology using an adaptive output-feedback approach,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 8, pp. 4676–4685, 2019.
- [9] K. Aryankia and R. R. Selmic, “Neural network-based formation control with target tracking for second-order nonlinear multiagent systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 1, pp. 328–341, 2022.
- [10] L. Cao, H. Li, G. Dong, and R. Lu, “Event-triggered control for multiagent systems with sensor faults and input saturation,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 6, pp. 3855–3866, 2021.
- [11] L. Wang, D. Zhu, W. Pang, and Y. Zhang, “A survey of underwater search for multi-target using multi-auv: Task allocation, path planning, and formation control,” *Ocean Engineering*, vol. 278, p. 114393, 2023.

- [12] X. Liu, X. Wu, Y. Xie, B. Guo, and J. Yan, “Fully distributed adaptive time-varying formation control of singular multiagent systems,” *Neurocomputing*, vol. 516, pp. 146–154, 2023.
- [13] L. Tian, Y. Hua, X. Dong, J. Lü, and Z. Ren, “Distributed time-varying group formation tracking for multiagent systems with switching interaction topologies via adaptive control protocols,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8422–8433, 2022.
- [14] E. Zhang, J. Li, C. Zhu, and B. Huang, “Bearing-based prescribed-time formation control of underactuated autonomous surface vehicles with the interception of attackers,” *Ocean Engineering*, vol. 262, p. 112187, 2022.
- [15] S. Zhao and D. Zelazo, “Translational and scaling formation maneuver control via a bearing-based approach,” *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 429–438, 2017.
- [16] X. Wang, B. Zerr, H. Thomas, B. Clement, and Z. Xie, “Pattern formation of multi-auv systems with the optical sensor based on displacement-based formation control,” *International Journal of Systems Science*, vol. 51, no. 2, pp. 348–367, 2020.
- [17] R. Babazadeh and R. R. Selmic, “Distance-based formation control of nonlinear agents over planar directed graphs,” in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 2321–2326.

- [18] F. Mehdifar, C. P. Bechlioulis, F. Hashemzadeh, and M. Baradarannia, “Prescribed performance distance-based formation control of multi-agent systems,” *Automatica*, vol. 119, p. 109086, 2020.
- [19] H. G. de Marina, “Maneuvering and robustness issues in undirected displacement-consensus-based formation control,” *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3370–3377, 2020.
- [20] Z. Lin, L. Wang, Z. Chen, M. Fu, and Z. Han, “Necessary and sufficient graphical conditions for affine formation control,” *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2877–2891, 2015.
- [21] S. Zhao, “Affine formation maneuver control of multiagent systems,” *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4140–4155, 2018.
- [22] S. M. Elkhider, S. El-Ferik, and A.-W. A. Saif, “Denial of service attack of qos-based control of multi-agent systems,” *Applied Sciences*, vol. 12, no. 9, p. 4315, 2022.
- [23] Y. Wang, L. Han, X. Li, X. Dong, Q. Li, and Z. Ren, “Resilient time-varying formation control of second-order discrete-time multi-agent systems with actuator faults and attacks on communication link,” in *Proceedings of 2021 5th Chinese Conference on Swarm Intelligence and Cooperative Control*. Springer, 2022, pp. 470–482.
- [24] S. Wang, X. Yin, P. Li, Y. Zhang, and H. Wen, “Consensus control of multi-agent systems with deception attacks using event-triggered adaptive cognitive

- control,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 114, p. 106675, 2022.
- [25] W. He, Z. Mo, Q.-L. Han, and F. Qian, “Secure impulsive synchronization in lipschitz-type multi-agent systems subject to deception attacks,” *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1326–1334, 2020.
- [26] W. Li, H. Zhang, W. Wang, and Z. Cao, “Fully distributed event-triggered time-varying formation control of multi-agent systems subject to mode-switching denial-of-service attacks,” *Applied Mathematics and Computation*, vol. 414, p. 126645, 2022.
- [27] T. Shi and F. Zhu, “Security time-varying formation control for multi-agent systems under denial-of-service attacks via unknown input observer,” *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 4, pp. 2372–2385, 2023.
- [28] L. Wang and Z. Wu, “Distributed secure consensus for first-order multi-agent systems under replay attacks,” in *2021 International Conference on Computer, Control and Robotics (ICCCR)*. IEEE, 2021, pp. 250–255.
- [29] M. Zhu and S. Martinez, “On the performance analysis of resilient networked control systems under replay attacks,” *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 804–808, 2013.

- [30] C. Zhang and Y. Lu, “Study on artificial intelligence: The state of the art and future prospects,” *Journal of Industrial Information Integration*, vol. 23, p. 100224, 2021.
- [31] P. Radanliev, D. De Roure, M. Van Kleek, O. Santos, and U. Ani, “Artificial intelligence in cyber physical systems,” *AI & society*, vol. 36, pp. 783–796, 2021.
- [32] Z. Lv, D. Chen, R. Lou, and A. Alazab, “Artificial intelligence for securing industrial-based cyber–physical systems,” *Future generation computer systems*, vol. 117, pp. 291–298, 2021.
- [33] Q.-L. Han, D. Ding, and X. Ge, “New trends of artificial-intelligence-based control, filtering, and optimization for industrial cyber-physical systems,” *Information Sciences*, p. 119333, 2023.
- [34] Z. Cuan, D.-W. Ding, Y. Yang, and Y. Xia, “Adaptive neural network finite-time control for nonlinear cyber-physical systems with external disturbances under malicious attacks,” *Neurocomputing*, vol. 518, pp. 133–141, 2023.
- [35] F. Farivar, M. S. Haghighi, A. Jolfaei, and M. Alazab, “Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial iot,” *IEEE transactions on industrial informatics*, vol. 16, no. 4, pp. 2716–2725, 2019.
- [36] Y. Yang, Y. Xiao, and T. Li, “A survey of autonomous underwater vehicle formation: Performance, formation control, and communication capability,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 815–841, 2021.

- [37] X. Li and L. Xie, “Dynamic formation control over directed networks using graphical laplacian approach,” *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3761–3774, 2018.
- [38] O. Onuoha, H. Thunay, and Z. Ding, “Affine formation maneuver control of multi-agent systems with triple-integrator dynamics,” in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5334–5339.
- [39] O. Onuoha, H. Thunay, Z. Li, and Z. Ding, “Affine formation algorithms and implementation based on triple-integrator dynamics,” *Unmanned Systems*, vol. 7, no. 01, pp. 33–45, 2019.
- [40] J. Wang, X. Ding, C. Wang, L. Liang, and H. Hu, “Affine formation control for multi-agent systems with prescribed convergence time,” *Journal of the Franklin Institute*, vol. 358, no. 14, pp. 7055–7072, 2021.
- [41] K. Gao, Y. Zhou, and Y. Zhao, “Affine formation control for multi-agent systems within an appointed settling time,” in *2021 40th Chinese Control Conference (CCC)*, 2021, pp. 5356–5361.
- [42] K. Gao, Y. Liu, Y. Zhou, Y. Zhao, and P. Huang, “Practical fixed-time affine formation for multi-agent systems with time-based generators,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 11, pp. 4433–4437, 2022.

- [43] F. Xiao, Q. Yang, X. Zhao, and H. Fang, “A framework for optimized topology design and leader selection in affine formation control,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8627–8634, 2022.
- [44] Z. Luo, P. Zhang, X. Ding, Z. Tang, C. Wang, and J. Wang, “Adaptive affine formation maneuver control of second-order multi-agent systems with disturbances,” in *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2020, pp. 1071–1076.
- [45] Y. Xu, D. Li, D. Luo, and Y. You, “Affine formation maneuver tracking control of multiple second-order agents with time-varying delays,” *Science China Technological Sciences*, vol. 62, pp. 665–676, 2019.
- [46] P. Zhang, D. Ma, P. Liu, and M. Li, “Robust affine maneuver formation control based of second-order multi-agent grid inspection systems,” *Frontiers in Energy Research*, vol. 10, p. 972999, 2022.
- [47] Y. Zhao, K. Gao, P. Huang, and G. Chen, “Specified-time affine formation maneuver control of multi-agent systems over directed networks,” *IEEE Transactions on Automatic Control*, pp. 1–8, 2023.
- [48] W. Yu, P. Yi, B. Zhu, H. Liu, and T. Hu, “Enhanced affine formation maneuver control using historical acceleration command,” *International Journal of Robust and Nonlinear Control*, pp. 1–23, 2023.

- [49] Y. Xu, S. Zhao, D. Luo, and Y. You, “Affine formation maneuver control of high-order multi-agent systems over directed networks,” *Automatica*, vol. 118, p. 109004, 2020.
- [50] L. Chen, J. Mei, C. Li, and G. Ma, “Distributed leader–follower affine formation maneuver control for high-order multiagent systems,” *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4941–4948, 2020.
- [51] X. Fang, X. Li, and L. Xie, “Distributed formation maneuver control of multiagent systems over directed graphs,” *IEEE Transactions on Cybernetics*, vol. 52, no. 8, pp. 8201–8212, 2022.
- [52] W. Yu, P. Yi, B. Zhu, H. Liu, and T. Hu, “Enhanced affine formation maneuver control using historical acceleration command,” *International Journal of Robust and Nonlinear Control*, vol. 4, no. 10, pp. 1–23.
- [53] W. Yu, B. Zhu, X. Wang, P. Yi, H. Liu, and T. Hu, “Enhanced affine formation maneuver control using historical velocity command (hvc),” *IEEE Robotics and Automation Letters*, vol. 8, no. 11, pp. 7186–7193, 2023.
- [54] X. Zhang, Q. Gao, and J. Cai, “Exact stability analysis of the second-order affine formation maneuver system with three independent time delays,” in *2023 42nd Chinese Control Conference (CCC)*, 2023, pp. 5529–5534.
- [55] Z. Li and R. T. Rajan, “Geometry-aware distributed kalman filtering for affine formation control under observation losses,” in *2023 26th International Conference on Information Fusion (FUSION)*, 2023, pp. 1–7.

- [56] X. Fang and L. Xie, “Distributed semi-global output feedback formation maneuver control of high-order multiagent systems,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 7, pp. 4498–4510, 2023.
- [57] Y. Xu, S. Zhao, D. Luo, and Y. You, “Affine formation maneuver control of multi-agent systems with directed interaction graphs,” in *2018 37th Chinese Control Conference (CCC)*. IEEE, 2018, pp. 4563–4568.
- [58] C. Wan, G. Jing, R. Dai, and R. Zhao, “Local shape-preserving formation maneuver control of multi-agent systems: From 2d to 3d,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 6251–6257.
- [59] O. Onuoha, H. Thunay, C. Wang, and Z. Ding, “Fully distributed affine formation control of general linear systems with uncertainty,” *Journal of the Franklin Institute*, vol. 357, no. 17, pp. 12 143–12 162, 2020.
- [60] O. Onuoha, H. Thunay, Z. Li, and Z. Ding, “Optimal affine formation control of linear multi-agent system,” in *2019 IEEE 15th International Conference on Control and Automation (ICCA)*. IEEE, 2019, pp. 851–856.
- [61] Y.-W. Chen, M.-L. Chiang, and L.-C. Fu, “Ordered formation control and affine transformation of multi-agent systems without global reference frame,” in *2019 American Control Conference (ACC)*, 2019, pp. 45–50.
- [62] Y. Xu, S. Zhao, D. Luo, and Y. You, “Affine formation maneuver control of linear multi-agent systems with undirected interaction graphs,” in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 502–507.

- [63] J. Wang, X. Ding, C. Wang, Z. Zuo, and Z. Ding, “Affine formation control of general linear multi-agent systems with delays,” *Unmanned Systems*, vol. 11, no. 02, pp. 123–132, 2023.
- [64] Z. Chang, W. Song, J. Wang, and Z. Li, “Fully distributed event-triggered affine formation maneuver control over directed graphs,” *Science China Information Sciences*, vol. 66, no. 6, pp. 1–3, 2023.
- [65] H. Zhi, L. Chen, C. Li, and Y. Lv, “Optimal leader-follower affine formation control of linear multi-agent systems,” *Optimal Control Applications and Methods*, vol. 43, no. 1, pp. 304–320, 2022.
- [66] Z. Chang, J. Wang, and Z. Li, “Fully distributed event-triggered affine formation maneuver control over directed graphs,” *IFAC-PapersOnLine*, vol. 55, no. 3, pp. 178–183, 2022.
- [67] Y. Xu, D. Luo, Y. You, and H. Duan, “Distributed adaptive affine formation control for heterogeneous linear networked systems,” *IEEE Access*, vol. 7, pp. 23 354–23 364, 2019.
- [68] H. Zhi and L. Chen, “Leader-follower affine formation control of second-order nonlinear uncertain multi-agent systems,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 12, pp. 3547–3551, 2021.
- [69] C. Zhu, B. Huang, Y. Xu, Y. Lu, and Y. Su, “Completely distributed affine formation maneuvering of networked marine surface vehicles with cooperation localization,” *IEEE Transactions on Vehicular Technology*, 2022.

- [70] Q. Yang, H. Fang, M. Cao, and J. Chen, “Planar affine formation stabilization via parameter estimations,” *IEEE Transactions on Cybernetics*, vol. 52, no. 6, pp. 5322–5332, 2020.
- [71] C. Zhu, B. Huang, Y. Lu, X. Li, and Y. Su, “Distributed affine formation maneuver control of autonomous surface vehicles with event-triggered data transmission mechanism,” *IEEE Transactions on Control Systems Technology*, pp. 1–12, 2022.
- [72] B. Zhou, B. Huang, L. Mao, C. Zhu, and Y. Su, “Distributed observer based event-triggered affine formation maneuver control for underactuated surface vessels with positive minimum inter-event times,” *International Journal of Robust and Nonlinear Control*, vol. 32, no. 14, pp. 7712–7732, 2022.
- [73] Y. Zheng, L. Zhang, B. Huang, and Y. Su, “Distributed event-triggered affine formation control for multiple underactuated marine surface vehicles,” *Ocean Engineering*, vol. 265, p. 112607, 2022.
- [74] L. Chen, Z. Sun, C. Li, B. Zhu, and C. Wang, “Satellite affine formation flying with obstacle avoidance,” *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 233, no. 16, pp. 5992–6004, 2019.
- [75] Y. Liu, X. Lin, and C. Zhang, “Affine formation maneuver control for multi-heterogeneous unmanned surface vessels in narrow channel environments,” *Journal of Marine Science and Engineering*, vol. 11, no. 9, p. 1811, 2023.

- [76] W. Pang, D. Zhu, C. Liu, and L. Wang, “The multi-auv time-varying formation reconfiguration control based on rigid-graph theory and affine transformation,” *Ocean Engineering*, vol. 270, p. 113521, 2023.
- [77] Y. Xu, W. Zheng, D. Luo, and H. Duan, “Dynamic affine formation control of networked under-actuated quad-rotor uavs with three-dimensional patterns,” *Journal of Systems Engineering and Electronics*, vol. 33, no. 6, pp. 1269–1285, 2022.
- [78] D. Li, G. Ma, Y. Xu, W. He, and S. S. Ge, “Layered affine formation control of networked uncertain systems: A fully distributed approach over directed graphs,” *IEEE Transactions on Cybernetics*, vol. 51, no. 12, pp. 6119–6130, 2021.
- [79] Y. Xu, D. Li, D. Luo, Y. You, and H. Duan, “Two-layer distributed hybrid affine formation control of networked euler–lagrange systems,” *Journal of the Franklin Institute*, vol. 356, no. 4, pp. 2172–2197, 2019.
- [80] H. Li, H. Chen, and X. Wang, “Affine formation tracking control of unmanned aerial vehicles,” *Frontiers of Information Technology & Electronic Engineering*, vol. 23, no. 6, pp. 909–919, 2022.
- [81] J. Li, Y. Fang, S. Huangfu, M. Zeng, G. Yang, and T. Deng, “Affine formation obstacle avoidance control of unmanned aerial vehicles with prescribed convergence time,” in *International Conference on Guidance, Navigation and Control*. Springer, 2022, pp. 4467–4478.

- [82] J. Yang, F. Xiao, and T. Chen, “Formation tracking of nonholonomic systems on the special euclidean group under fixed and switching topologies: an affine formation strategy,” *SIAM Journal on Control and Optimization*, vol. 59, no. 4, pp. 2850–2874, 2021.
- [83] W. He, W. Xu, X. Ge, Q.-L. Han, W. Du, and F. Qian, “Secure control of multi-agent systems against malicious attacks: A brief survey,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 3595–3608, 2021.
- [84] S. Xiao, X. Ge, Q.-L. Han, and Y. Zhang, “Secure and collision-free multi-platoon control of automated vehicles under data falsification attacks,” *Automatica*, vol. 145, p. 110531, 2022.
- [85] M. Xing, J. Lu, Y. Liu, and X. Chen, “Event-based bipartite consensus of multi-agent systems subject to dos attacks,” *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 1, pp. 68–80, 2022.
- [86] L. Zhou, X. Chen, and H. Wen, “Event-triggered h_∞ consensus for nonlinear multi-agent systems with semi-markov switching topologies under dos attacks,” *Franklin Open*, vol. 2, p. 100006, 2023.
- [87] D. Zhang, C. Deng, and G. Feng, “Resilient cooperative output regulation for nonlinear multi-agent systems under dos attacks,” *IEEE Transactions on Automatic Control*, pp. 1–8, 2022.
- [88] G. Narayanan, M. S. Ali, Q. Zhu, B. Priya, and G. K. Thakur, “Fuzzy observer-based consensus tracking control for fractional-order multi-agent systems under

- cyber-attacks and its application to electronic circuits,” *IEEE Transactions on Network Science and Engineering*, pp. 1–11, 2022.
- [89] A. Mousavi, K. Aryankia, and R. R. Selmic, “A distributed fdi cyber-attack detection in discrete-time nonlinear multi-agent systems using neural networks,” *European Journal of Control*, vol. 66, p. 100646, 2022.
- [90] S. Wang, S. Zheng, C. Zhao, H. Jian, and H. Li, “Formation control of nonlinear multi-agent systems with actuator and communication attacks,” in *2021 40th Chinese Control Conference (CCC)*, 2021, pp. 2286–2291.
- [91] L. Xia, Q. Li, R. Song, and Z. Zhang, “Leader-follower time-varying output formation control of heterogeneous systems under cyber attack with active leader,” *Information Sciences*, vol. 585, pp. 24–40, 2022.
- [92] Y. Wang, L. Han, X. Li, X. Dong, Q. Li, and Z. Ren, “Resilient time-varying formation control of second-order discrete-time multi-agent systems with actuator faults and attacks on communication link,” in *Proceedings of 2021 5th Chinese Conference on Swarm Intelligence and Cooperative Control*. Springer, 2023, pp. 470–482.
- [93] Y. Yang, Y. Xiao, and T. Li, “Attacks on formation control for multiagent systems,” *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 805–12 817, 2022.

- [94] Y. Tang, D. Zhang, P. Shi, W. Zhang, and F. Qian, “Event-based formation control for nonlinear multiagent systems under dos attacks,” *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 452–459, 2020.
- [95] S. M. Elkhider, S. El-Ferik, and A.-W. A. Saif, “Containment control of multiagent systems subject to denial of service attacks,” *IEEE Access*, vol. 10, pp. 48 102–48 111, 2022.
- [96] K. Pan, Y. Lyu, and Q. Pan, “Adaptive formation for multiagent systems subject to denial-of-service attacks,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 8, pp. 3391–3401, 2022.
- [97] J. Gong, B. Jiang, Y. Ma, and Z. Mao, “Distributed adaptive fault-tolerant formation control for heterogeneous multiagent systems with communication link faults,” *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–11, 2022.
- [98] Y. Ouyang, C. Sun, and L. Dong, “Actor–critic learning based coordinated control for a dual-arm robot with prescribed performance and unknown backlash-like hysteresis,” *ISA transactions*, vol. 126, pp. 1–13, 2022.
- [99] L. Xi, J. Wu, Y. Xu, and H. Sun, “Automatic generation control based on multiple neural networks with actor-critic strategy,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 6, pp. 2483–2493, 2020.
- [100] X. Wang, Q. Wang, and C. Sun, “Prescribed performance fault-tolerant control for uncertain nonlinear mimo system using actor–critic learning structure,”

- IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 9, pp. 4479–4490, 2022.
- [101] Y. Sun, J. Xu, H. Wu, G. Lin, and S. Mumtaz, “Deep learning based semi-supervised control for vertical security of maglev vehicle with guaranteed bounded airgap,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4431–4442, 2021.
- [102] D. Hu, Z. Ye, Y. Gao, Z. Ye, Y. Peng, and N. Yu, “Multi-agent deep reinforcement learning for voltage control with coordinated active and reactive power optimization,” *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4873–4886, 2022.
- [103] Y. Deng, T. Liu, and D. Zhao, “Event-triggered output-feedback adaptive tracking control of autonomous underwater vehicles using reinforcement learning,” *Applied Ocean Research*, vol. 113, p. 102676, 2021.
- [104] Z. Peng, Y. Zhao, J. Hu, R. Luo, B. K. Ghosh, and S. K. Nguang, “Input–output data-based output antisynchronization control of multiagent systems using reinforcement learning approach,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7359–7367, 2021.
- [105] Z. Peng, R. Luo, J. Hu, K. Shi, S. K. Nguang, and B. K. Ghosh, “Optimal tracking control of nonlinear multiagent systems using internal reinforce q-learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 4043–4055, 2022.

- [106] G. Wen and B. Li, “Optimized leader-follower consensus control using reinforcement learning for a class of second-order nonlinear multiagent systems,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 9, pp. 5546–5555, 2021.
- [107] L. Zhang, J. Li, H. Shi, K.-S. Hwang *et al.*, “Multi-agent reinforcement learning by the actor-critic model with an attention interface,” *Neurocomputing*, vol. 471, pp. 275–284, 2022.
- [108] H. Liu, F. Peng, H. Modares, and B. Kiumarsi, “Heterogeneous formation control of multiple rotorcrafts with unknown dynamics by reinforcement learning,” *Information Sciences*, vol. 558, pp. 194–207, 2021.
- [109] Y. Fei, P. Shi, and C.-C. Lim, “Robust formation control for multi-agent systems: A reference correction based approach,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 6, pp. 2616–2625, 2021.
- [110] W. Zhao, H. Liu, K. P. Valavanis, and F. L. Lewis, “Fault-tolerant formation control for heterogeneous vehicles via reinforcement learning,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 4, pp. 2796–2806, 2021.
- [111] H. Liu, W. Zhao, and J. Xi, “Optimal formation control for a quadrotor team under switching topologies via reinforcement learning,” in *2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*, 2021, pp. 731–736.

- [112] W. Zhao, H. Liu, and F. L. Lewis, “Robust formation control for cooperative underactuated quadrotors via reinforcement learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 10, pp. 4577–4587, 2020.
- [113] Y. Yin, F. Wang, Z. Liu, and Z. Chen, “Finite-time leader-following consensus of multiagent systems with actuator faults and input saturation,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 5, pp. 3314–3325, 2022.
- [114] Y. Zhao, X. Du, C. Zhou, Y.-C. Tian, X. Hu, and D. E. Quevedo, “Adaptive resilient control of cyber-physical systems under actuator and sensor attacks,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3203–3212, 2022.
- [115] B. Chen, Y. Tan, Z. Sun, and L. Yu, “Attack-resilient control against fdi attacks in cyber-physical systems,” *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 6, pp. 1099–1102, 2022.
- [116] Z. Cao, Y. Niu, and Y. Zou, “Adaptive neural sliding mode control for singular semi-markovian jump systems against actuator attacks,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1523–1533, 2021.
- [117] Z. Jing, Y. Ma, X. Wu, X. He, and Y. Sun, “Backstepping control for vibration suppression of 2-d euler–bernoulli beam based on nonlinear saturation compensator,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 5, pp. 2562–2571, 2023.

- [118] K. Zhao, Y. Song, C. P. Chen, and L. Chen, “Adaptive asymptotic tracking with global performance for nonlinear systems with unknown control directions,” *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1566–1573, 2021.
- [119] Y. Liu, S. Sheng, J. Zhang, and G. Lu, “Bipartite synchronization of multi-agent systems with signed graph under deception attacks and application in speech communication,” *IEEE Access*, vol. 11, pp. 15 954–15 967, 2022.
- [120] S. Wang, C. Zhao, B. Zhang, and Y. Jiang, “Event-triggered based security consensus control for multi-agent systems with dos attacks,” *Neurocomputing*, vol. 505, pp. 214–224, 2022.
- [121] X.-M. Li, W. Xiao, G. Lin, H. Li, and R. Lu, “Observer-based security control for distributed cyber-physical systems under replay attacks,” *International Journal of Robust and Nonlinear Control*, vol. 33, no. 13, pp. 8015–8035, 2023.
- [122] T. Yin, Z. Gu, and J. H. Park, “Event-based intermittent formation control of multi-uav systems under deception attacks,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2022.
- [123] B. Sun, Z. Gu, and T. Xiong, “Event-triggered formation tracking control for unmanned aerial vehicles subjected to deception attacks,” *Electronics*, vol. 10, no. 22, p. 2736, 2021.
- [124] Y. Zhang, L. Ma, C. Yang, and L. Zhou, “Formation control for quadrotor swarms with deception attacks via singular perturbation,” in *2022 IEEE Inter-*

- national Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2022, pp. 189–194.
- [125] B.-S. Park and S.-J. Yoo, “Adaptive secure control for leader-follower formation of nonholonomic mobile robots in the presence of uncertainty and deception attacks,” *Mathematics*, vol. 9, no. 18, p. 2190, 2021.
- [126] Y.-C. Liu, K.-Y. Liu, and Z. Song, “Resilient time-varying formation tracking for mobile robot networks under deception attacks on positioning,” *International Journal of Robust and Nonlinear Control*, vol. 33, no. 11, p. 6308–6328, 2023.
- [127] S. Liu, L. Zhang, B. Niu, X. Zhao, and A. M. Ahmad, “Adaptive neural finite-time hierarchical sliding mode control of uncertain under-actuated switched nonlinear systems with backlash-like hysteresis,” *Information Sciences*, vol. 599, pp. 147–169, 2022.
- [128] R. Dong, Y. Tan, and K. Janschek, “Nonsmooth predictive control for wiener systems with backlash-like hysteresis,” *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 1, pp. 17–28, 2015.
- [129] Y. Liu, H. Li, R. Lu, Z. Zuo, and X. Li, “An overview of finite/fixed-time control and its application in engineering systems,” *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2106–2120, 2022.
- [130] K. Liu, P. Yang, R. Wang, L. Jiao, T. Li, and J. Zhang, “Observer-based adaptive fuzzy finite-time attitude control for quadrotor uavs,” *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–17, 2023.

- [131] L. Ding, W. Wang, and Y. Yu, “Finite-time adaptive nn control for permanent magnet synchronous motors with full-state constraints,” *Neurocomputing*, vol. 449, pp. 435–442, 2021.
- [132] L. Liu, Z. Wang, and H. Zhang, “Adaptive fault-tolerant tracking control for mimo discrete-time systems via reinforcement learning algorithm with less learning parameters,” *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 299–313, 2016.
- [133] X. Wang, C. Zhao, T. Huang, P. Chakrabarti, and J. Kurths, “Cooperative learning of multi-agent systems via reinforcement learning,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 9, pp. 13–23, 2023.
- [134] X. Guo, W. Yan, and R. Cui, “Integral reinforcement learning-based adaptive nn control for continuous-time nonlinear mimo systems with unknown control directions,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 11, pp. 4068–4077, 2019.
- [135] Y. Yang, J. Huang, X. Su, K. Wang, and G. Li, “Adaptive control of second-order nonlinear systems with injection and deception attacks,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 1, pp. 574–581, 2020.
- [136] Q. Wan, W.-H. Chen, and X. Lu, “Secure consensus tracking of multi-agent systems with network-induced delays under deception attacks via guaranteed performance impulsive control,” *Nonlinear Dynamics*, pp. 1–20, 2023.

- [137] D. Zhang, G. Feng, Y. Shi, and D. Srinivasan, “Physical safety and cyber security analysis of multi-agent systems: A survey of recent advances,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 319–333, 2021.
- [138] H. Modares, R. Moghadam, F. L. Lewis, and A. Davoudi, “Static output-feedback synchronisation of multi-agent systems: a secure and unified approach,” *IET Control Theory & Applications*, vol. 12, no. 8, pp. 1095–1106, 2018.
- [139] Z. Li and J. Zhao, “Resilient adaptive control of switched nonlinear cyber-physical systems under uncertain deception attacks,” *Information sciences*, vol. 543, pp. 398–409, 2021.
- [140] Z. Han, W. Wang, J. Huang, and Z. Wang, “Distributed adaptive formation tracking control of mobile robots with event-triggered communication and denial-of-service attacks,” *IEEE Transactions on Industrial Electronics*, vol. 70, no. 4, pp. 4077–4087, 2022.
- [141] Y. Liu, X. Dong, Z. Ren, and J. Liu, “Secure formation control for resilient multi-agent cyber-physical systems under dos attacks and faults,” *International Journal of Robust and Nonlinear Control*, vol. 33, no. 6, pp. 3607–3626, 2023.
- [142] Y. Zhang, L. Ma, C. Yang, L. Zhou, G. Wang, and W. Dai, “Formation control for multiple quadrotors under dos attacks via singular perturbation,” *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–10, 2023.

- [143] W. Zhan, Z. Miao, Y. Chen, Z.-G. Wu, and Y. Wang, “Event-triggered finite-time formation control for networked nonholonomic mobile robots under denial-of-service attacks,” *IEEE Transactions on Network Science and Engineering*, pp. 1–13, 2023.
- [144] W. Wang, Z. Han, K. Liu, and J. Lü, “Distributed adaptive resilient formation control of uncertain nonholonomic mobile robots under deception attacks,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 9, pp. 3822–3835, 2021.
- [145] Z. Yu, Y. Zhang, B. Jiang, C.-Y. Su, J. Fu, Y. Jin, and T. Chai, “Decentralized fractional-order backstepping fault-tolerant control of multi-uavs against actuator faults and wind effects,” *Aerospace Science and Technology*, vol. 104, pp. 105 939–57, 2020.
- [146] W. Cheng, K. Zhang, and B. Jiang, “Continuous fixed-time fault-tolerant formation control for heterogeneous multiagent systems under fixed and switching topologies,” *IEEE Transactions on Vehicular Technology*, vol. 72, no. 2, pp. 1545–1558, 2023.
- [147] H. Wang and J. Dong, “Time-varying formation control for multi-agent systems under directed topology base on gain re-adaptation fault-tolerant compensation approach,” *International Journal of Robust and Nonlinear Control*, vol. 32, no. 6, pp. 3909–3922, 2022.

- [148] H. Zhou, S. Sui, and S. Tong, “Finite-time adaptive fuzzy prescribed performance formation control for high-order nonlinear multiagent systems based on event-triggered mechanism,” *IEEE Transactions on Fuzzy Systems*, vol. 31, no. 4, pp. 1229–1239, 2023.
- [149] F. Chen and D. V. Dimarogonas, “Leader–follower formation control with prescribed performance guarantees,” *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 450–461, 2020.
- [150] Y. Yang, X. Si, D. Yue, and Y.-C. Tian, “Time-varying formation tracking with prescribed performance for uncertain nonaffine nonlinear multiagent systems,” *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 4, pp. 1778–1789, 2021.
- [151] H. Su and G.-Y. Tang, “Rolling optimization formation control for multi-agent systems under unknown prior desired shapes,” *Information Sciences*, vol. 459, pp. 255–264, 2018.
- [152] Y. Liu and Z. Geng, “Finite-time optimal formation tracking control of vehicles in horizontal plane,” *Nonlinear Dynamics*, vol. 76, pp. 481–495, 2014.
- [153] H. Liu, Z. Chen, X. Wang, and Z. Sun, “Optimal formation control for multiple rotation-translation coupled satellites using reinforcement learning,” *Acta Astronautica*, vol. 204, pp. 583–590, 2023.
- [154] W. Shao, Y. Chen, and J. Huang, “Optimized formation control for a class of second-order multi-agent systems based on single critic reinforcement learning

- method,” in *2021 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, vol. 1. IEEE, 2021, pp. 1–6.
- [155] G. Wen, C. P. Chen, and S. S. Ge, “Simplified optimized backstepping control for a class of nonlinear strict-feedback systems with unknown dynamic functions,” *IEEE Transactions on Cybernetics*, vol. 51, no. 9, pp. 4567–4580, 2020.
- [156] J. Zhang, K. Li, and Y. Li, “Output-feedback based simplified optimized backstepping control for strict-feedback systems with input and state constraints,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 6, pp. 1119–1132, 2021.
- [157] Y. Zhang, X. Liang, D. Li, S. S. Ge, B. Gao, H. Chen, and T. H. Lee, “Barrier lyapunov function-based safe reinforcement learning for autonomous vehicles with optimized backstepping,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2022.
- [158] X. Yang and B. Li, “Optimized tracking control using reinforcement learning strategy for a class of nonlinear systems,” *Asian Journal of Control*, vol. 25, no. 3, pp. 2095–2104, 2023.
- [159] Y. Li, Y. Liu, and S. Tong, “Observer-based neuro-adaptive optimized control of strict-feedback nonlinear systems with state constraints,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 7, pp. 3131–3145, 2022.
- [160] G. Wen and C. L. P. Chen, “Optimized backstepping consensus control using reinforcement learning for a class of nonlinear strict-feedback-dynamic multi-

- agent systems,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 3, pp. 1524–1536, 2023.
- [161] J. Huang, Z. Zhang, F. Cai, and Y. Chen, “Optimized formation control for multi-agent systems based on adaptive dynamic programming without persistence of excitation,” *IEEE Control Systems Letters*, vol. 6, pp. 1412–1417, 2021.
- [162] J. Lan, Y.-J. Liu, D. Yu, G. Wen, S. Tong, and L. Liu, “Time-varying optimal formation control for second-order multiagent systems based on neural network observer and reinforcement learning,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2022.
- [163] H. Zhi, L. Chen, C. Li, and Y. Guo, “Leader–follower affine formation control of second-order nonlinear uncertain multi-agent systems,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 12, pp. 3547–3551, 2021.
- [164] Y. Xu, D. Luo, Y. You, and H. Duan, “Affine transformation based formation maneuvering for discrete-time directed networked systems,” *Science China Technological Sciences*, vol. 63, no. 1, pp. 73–85, 2020.
- [165] X. Chai, Y. Sun, N. Zhang, S. Ma, Y. Wang, and Q. Diao, “Sampled-data-based event-triggered affine formation control for multi-agent systems with integrator dynamics,” in *2022 IEEE International Conference on Unmanned Systems (ICUS)*, 2022, pp. 715–719.

- [166] Z. Wu, J. Ni, W. Qian, X. Bu, and B. Liu, “Composite prescribed performance control of small unmanned aerial vehicles using modified nonlinear disturbance observer,” *ISA transactions*, vol. 116, pp. 30–45, 2021.
- [167] W. Liu, S. Fei, Q. Ma, H. Zhao, and S. Xu, “Prescribed performance dynamic surface control for nonlinear systems subject to partial and full state constraints,” *Applied Mathematics and Computation*, vol. 431, p. 127318, 2022.
- [168] Y. Zhou and S. Wang, “Adaptive prescribed performance control for nonlinear robotic systems,” *Journal of the Franklin Institute*, vol. 360, no. 2, pp. 1378–1394, 2023.
- [169] J. Du, J. Li, and F. L. Lewis, “Distributed 3-d time-varying formation control of underactuated auvs with communication delays based on data-driven state predictor,” *IEEE Transactions on Industrial Informatics*, vol. 19, no. 5, pp. 6963–6971, 2023.
- [170] Y. Wang, Y. Liu, Y. Wang, and T. Chai, “Neural output feedback control of automobile steer-by-wire system with predefined performance and composite learning,” *IEEE Transactions on Vehicular Technology*, vol. 72, no. 5, pp. 5906–5921, 2023.
- [171] Y. Liu, X. Liu, and Y. Jing, “Adaptive neural networks finite-time tracking control for non-strict feedback systems via prescribed performance,” *Information Sciences*, vol. 468, pp. 29–46, 2018.

- [172] Z. Gao, Y. Zhang, and G. Guo, “Fixed-time prescribed performance adaptive fixed-time sliding mode control for vehicular platoons with actuator saturation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 24 176–24 189, 2022.
- [173] Y. Yang, D. Chen, Q. Liu, T. Zhang, A. Liu, and W. Yue, “Predictor-based neural dynamic surface control of a nontriangular system with unknown disturbances,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 8, pp. 3353–3365, 2022.
- [174] M. Shahvali, M.-B. Naghibi-Sistani, and J. Askari, “Adaptive output-feedback bipartite consensus for nonstrict-feedback nonlinear multi-agent systems: A finite-time approach,” *Neurocomputing*, vol. 318, pp. 7–17, 2018.
- [175] T. Yin, Z. Gu, and S. Yan, “Event-based formation control for multiple unmanned aerial vehicles under directed topology,” *ISA transactions*, vol. 137, pp. 111–121, 2023.
- [176] J. A. Vazquez Trejo, A. Guenard, M. Adam-Medina, L. Ciarletta, J.-C. Ponsart, and D. Theilliol, “Resilient leader-following formation control for a fleet of unmanned aerial vehicles under cyber-attacks,” in *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2023, pp. 731–737.
- [177] Z. Hu and X. Jin, “Adaptive formation control architectures for a team of quadrotors with multiple performance and safety constraints,” *International Journal of Robust and Nonlinear Control*, vol. n/a, no. n/a, pp. 1–22, 2023.

- [178] B. An, B. Wang, H. Fan, L. Liu, H. Hu, and Y. Wang, “Fully distributed prescribed performance formation control for uavs with unknown maneuver of leader,” *Aerospace Science and Technology*, vol. 130, p. 107886, 2022.
- [179] F. Wang, Y. Gao, C. Zhou, and Q. Zong, “Disturbance observer-based backstepping formation control of multiple quadrotors with asymmetric output error constraints,” *Applied Mathematics and Computation*, vol. 415, p. 126693, 2022.
- [180] W. Zhao, H. Liu, and F. L. Lewis, “Data-driven fault-tolerant control for attitude synchronization of nonlinear quadrotors,” *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5584–5591, 2021.
- [181] W. Zhao, H. Liu, and Y. Wan, “Data-driven fault-tolerant formation control for nonlinear quadrotors under multiple simultaneous actuator faults,” *Systems & Control Letters*, vol. 158, p. 105063, 2021.
- [182] H. Liu, Q. Meng, F. Peng, and F. L. Lewis, “Heterogeneous formation control of multiple uavs with limited-input leader via reinforcement learning,” *Neurocomputing*, vol. 412, pp. 63–71, 2020.
- [183] H. Liu, F. Peng, H. Modares, and B. Kiumarsi, “Heterogeneous formation control of multiple rotorcrafts with unknown dynamics by reinforcement learning,” *Information Sciences*, vol. 558, pp. 194–207, 2021.
- [184] W. Zhao, H. Liu, J. Lü, Q. Gao, and G. Feng, “Data-driven optimal formation control for multiple nonlinear quadrotors with switching topologies,” *IEEE Transactions on Vehicular Technology*, 2023.

- [185] Y. Sun, J. Li, Z. Wang, X. He, Q. Fu, and Y. Zou, “Distributed formation-aggregation control algorithm for a cluster of quadrotors,” *Journal of the Franklin Institute*, vol. 360, no. 3, pp. 1560–1581, 2023.
- [186] W. Yu, P. Yi, B. Zhu, H. Liu, and T. Hu, “Enhanced affine formation maneuver control using historical acceleration command,” *International Journal of Robust and Nonlinear Control*, pp. 1–23, 2023.
- [187] S. El Ferik and A. Aliyu, “Adaptive backstepping multi-agent control of multiple uavs conveying load: A graph rigidity approach,” *IEEE Access*, vol. 11, pp. 58 269–58 278, 2023.
- [188] B. Tian, L. Liu, H. Lu, Z. Zuo, Q. Zong, and Y. Zhang, “Multivariable finite time attitude control for quadrotor uav: Theory and experimentation,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2567–2577, 2018.
- [189] X.-Z. Jin, W.-W. Che, Z.-G. Wu, and C. Deng, “Robust adaptive general formation control of a class of networked quadrotor aircraft,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 12, pp. 7714–7726, 2022.
- [190] X. Fu, J. Pan, H. Wang, and X. Gao, “A formation maintenance and reconstruction method of uav swarm based on distributed control,” *Aerospace Science and Technology*, vol. 104, p. 105981, 2020.

- [191] J. Hu, H. Niu, J. Carrasco, B. Lennox, and F. Arvin, “Fault-tolerant cooperative navigation of networked uav swarms for forest fire monitoring,” *Aerospace Science and Technology*, vol. 123, p. 107494, 2022.
- [192] J. González-Sierra, A. Dzul, and E. Martínez, “Formation control of distance and orientation based-model of an omnidirectional robot and a quadrotor uav,” *Robotics and Autonomous Systems*, vol. 147, p. 103921, 2022.
- [193] J. Xia, J. Zhang, J. Feng, Z. Wang, and G. Zhuang, “Command filter-based adaptive fuzzy control for nonlinear systems with unknown control directions,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1945–1953, 2019.
- [194] Y.-J. Liu and S. Tong, “Barrier lyapunov functions for nussbaum gain adaptive control of full state constrained nonlinear systems,” *Automatica*, vol. 76, pp. 143–152, 2017.
- [195] Y.-J. Liu, M. Gong, S. Tong, C. L. P. Chen, and D.-J. Li, “Adaptive fuzzy output feedback control for a class of nonlinear systems with full state constraints,” *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 2607–2617, 2018.
- [196] Y.-J. Liu, J. Li, S. Tong, and C. L. P. Chen, “Neural network control-based adaptive learning design for nonlinear systems with full-state constraints,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 7, pp. 1562–1571, 2016.

- [197] C. Garanayak and D. Mukherjee, “Leaderless affine formation maneuvers over directed graphs,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 3983–3988.
- [198] H. Han, J. Zhang, H. Yang, Y. Hou, and J. Qiao, “Data-driven robust optimal control for nonlinear system with uncertain disturbances,” *Information Sciences*, vol. 621, pp. 248–264, 2023.
- [199] A. Das, F. Lewis, and K. Subbarao, “Backstepping approach for controlling a quadrotor using lagrange form dynamics,” *Journal of Intelligent and Robotic Systems*, vol. 56, pp. 127–151, 2009.
- [200] J. Sun, Y. Wang, Y. Yu, and C. Sun, “Nonlinear robust compensation method for trajectory tracking control of quadrotors,” *IEEE Access*, vol. 7, pp. 26 766–26 776, 2019.
- [201] O. Onuoha, H. Tnunay, C. Wang, and Z. Ding, “Fully distributed affine formation control of general linear systems with uncertainty,” *Journal of the Franklin Institute*, vol. 357, no. 17, pp. 12 143–12 162, 2020.
- [202] W.-H. Chen, Q. Wan, and X. Lu, “Guaranteed performance impulsive tracking control of multi-agents systems under discrete-time deception attacks,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 117, p. 106905, 2023.
- [203] Z. Wang, S. Shi, W. He, M. Xiao, J. Cao, and S. Gorbachev, “Observer-based asynchronous event-triggered bipartite consensus of multi-agent systems under

- false data injection attacks,” *IEEE Transactions on Control of Network Systems*, pp. 1–12, 2023.
- [204] R. Nie, W. Du, Z. Li, and S. He, “Sliding mode-based finite-time consensus tracking control for multi-agent systems under actuator attacks,” *Information Sciences*, vol. 640, p. 118971, 2023.
- [205] Z. Chang, H. Xue, H. Liang, and P. Zhang, “Time-varying formation control for nonlinear multi-agent systems against actuator attacks,” *Journal of the Franklin Institute*, vol. 359, no. 18, pp. 11 068–11 088, 2022.
- [206] Y. Yang, Y. Xiao, and T. Li, “Attacks on formation control for multiagent systems,” *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 805–12 817, 2021.
- [207] S. El-Ferik, M. Maaruf, F. M. Al-Sunni, A. A. Saif, and M. M. Al Dhaifallah, “Reinforcement learning-based control strategy for multi-agent systems subjected to actuator cyberattacks during affine formation maneuvers,” *IEEE Access*, vol. 11, pp. 77 656–77 668, 2023.
- [208] K. G. Vamvoudakis and F. L. Lewis, “Online solution of nonlinear two-player zero-sum games using synchronous policy iteration,” *International Journal of Robust and Nonlinear Control*, vol. 22, no. 13, pp. 1460–1483, 2012.
- [209] G. Xiao, H. Zhang, K. Zhang, and Y. Wen, “Value iteration based integral reinforcement learning approach for h_∞ controller design of continuous-time nonlinear systems,” *Neurocomputing*, vol. 285, pp. 51–59, 2018.

- [210] G. Xiao, H. Zhang, Q. Qu, and H. Jiang, “General value iteration based single network approach for constrained optimal controller design of partially-unknown continuous-time nonlinear systems,” *Journal of the Franklin Institute*, vol. 355, no. 5, pp. 2610–2630, 2018.
- [211] D. Wang, H. He, C. Mu, and D. Liu, “Intelligent critic control with disturbance attenuation for affine dynamics including an application to a microgrid system,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4935–4944, 2017.
- [212] Y. Gao, C. Liu, S. Jiang, and S. Zhang, “Zero-sum differential games-based fast adaptive robust optimal sliding mode control design for uncertain missile autopilot with constrained input,” *International Journal of Control*, vol. 95, no. 7, pp. 1789–1801, 2022.
- [213] G. Ibragimov, M. Ruziboev, I. Zaynabiddinov, and B. A. Pansera, “Evasion differential game of multiple pursuers and a single evader with geometric constraints in l_2 ,” *Games*, vol. 14, no. 4, pp. 52–57, 2023.
- [214] Y. Yuan, P. Zhang, and X. Li, “Synchronous fault-tolerant near-optimal control for discrete-time nonlinear pe game,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 10, pp. 4432–4444, 2021.
- [215] C. Wu, L. Wu, J. Liu, and Z.-P. Jiang, “Active defense-based resilient sliding mode control under denial-of-service attacks,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 237–249, 2020.

- [216] Y. Zhou, K. G. Vamvoudakis, W. M. Haddad, and Z.-P. Jiang, “A secure control learning framework for cyber-physical systems under sensor attacks,” in *2019 American Control Conference (ACC)*, 2019, pp. 4280–4285.
- [217] H. Ren, B. Jiang, and Y. Ma, “Zero-sum differential game-based fault-tolerant control for a class of affine nonlinear systems,” *IEEE Transactions on Cybernetics*, pp. 1–11, 2022.
- [218] Y. Zhang, L. Ma, C. Yang, and W. Dai, “Reinforcement learning-based sliding mode tracking control for the two-time-scale systems: Dealing with actuator attacks,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 9, pp. 3819–3823, 2022.

VITAE

PERSONAL INFORMATION

- Name: Muhammad Maaruf
- Nationality: Nigerian
- Date of Birth: 18/Oct/1992
- Email: *mdanlamimaaruf@gmail.com*
- Permanent Address: VLC 283, Janbulo Kabuga Housing Estate, Kano, 700103, Nigeria.

EDUCATION BACKGROUND

- B. Eng. Electrical Engineering, Kano University of Science and Technology, Kano, Nigeria, 2016.
- M.Sc. in Systems and Control Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, 2020.
- Ph.D. in Systems and Control Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, 2023.

LIST of PUBLICATIONS

JOURNAL PAPERS

1. **M. Maaruf**, S. El-Ferik, A. A. Saif, F. M. Al-Sunni and M. M. Al-Dhaifallah
"Affine Formation Maneuvering of Nonlinear Multi-Agent Systems with Fault-Tolerant Secure Optimized Backstepping Control Using Reinforcement Learning", (Submitted to IEEE Transaction on Cybernetics).
[Scopus indexed; Journal Rank: Q1; Impact factor: 11.8](#)
[\(From my Ph.D. dissertation\)](#)
2. **M. Maaruf** and S. El-Ferik, "Learning-Based Safety-Guaranteed Sliding Mode Affine Formation Maneuver Control of Quadrotors Vulnerable to Cyber-Attacks", (Submitted to ISA Transactions).
[Scopus indexed; Journal Rank: Q1; Impact factor: 7.3](#)
[\(From my Ph.D. dissertation\)](#)
3. **M. Maaruf** and S. El-Ferik, "Critic Learning Control Via Zero-Sum Differential Game for Affine Formation Maneuver of Multi-agent Systems with Cyber-Attacks", (Submitted to International Journal of Systems Sciences).
[Scopus indexed; Journal Rank: Q1; Impact factor: 4.3](#)
[\(From my Ph.D. dissertation\)](#)
4. S. A. Kashif, T. Shami, M. A. Aslam, M. M. Gulzar, **M. Maaruf**, F. Alismaili nad M. Khalid, "Selective Harmonic Elimination Notch Angle Calculation Using THD and ZHF Benchmarks for Cascaded Multilevel Inverters," *in IEEE Access*, vol. 11, pp. 116497-116510, 2023, doi: 10.1109/ACCESS.2023.3324557.

Scopus indexed; Journal Rank: Q2; Impact factor: 3.9

5. S. El-Ferik, **M. Maaruf**, F. M. Al-Sunni, A. A. Saif and M. M. Al Dhaifallah, "Reinforcement Learning-Based Control Strategy for Multi-Agent Systems Subjected to Actuator Cyberattacks During Affine Formation Maneuvers," in *IEEE Access*, vol. 11, pp. 77656-77668, 2023, doi: 10.1109/ACCESS.2023.3296741.

Scopus indexed; Journal Rank: Q2; Impact factor: 3.9

(From my Ph.D. dissertation)

6. **M. Maaruf**, M.A. Abido, W.M. Hamanah. Hybrid Backstepping Control of a Quadrotor Using a Radial Basis Function Neural Network. *Mathematics*, 11(4), p.991. 2023

Scopus indexed; Journal Rank: Q1; Impact factor: 2.592

7. **M. Maaruf**, M.A.M. Eltoum, F.M. Al-Sunni. Artificial intelligence based control of continuous polymerization reactor with input dead-zone. *Int. J. Dynam. Control*, 1-13, 2022.

<https://link.springer.com/article/10.1007/s40435-022-01038-9>

Scopus indexed; Journal Rank: Q2; Impact factor: 2.016

8. **M. Maaruf** , A. Babangida, HA Almusawi, PS Tamas PS. Neural Network-based Finite-time Control of Nonlinear Systems with Unknown Dead-zones: Application to Quadrotors. *Journal of Robotics and Control (JRC)*, 3(6):735-42,

2022

<https://journal.umy.ac.id/index.php/jrc/article/view/15355>

Scopus indexed; Journal Rank: Q2; Impact factor: 2.738

9. **M. Maaruf**, M.S. Mahmoud, A. Ma'arif. A Survey of Control Methods for Quadrotor UAV. *International Journal of Robotics and Control Systems*, 2(4):652-65, 2022

<https://pubs2.ascee.org/index.php/IJRCS/article/view/743>

Scopus indexed; Journal Rank: Impact factor:

10. **M. Maaruf**, M.S. Mahmoud, S. El-Ferik. Adaptive Control of Continuous Polymerization Reactor. *Cogent Engineering*, 9(1), 2086673-2086688, 2022.

(From my master's thesis)

Scopus indexed; Journal Rank: Q2; Impact factor: 1.689

11. **M. Maaruf**, K. A. Khan, and M. Khalid. Robust Control for Optimized Islanded and Grid-Connected Operation of Solar/Wind/Battery Hybrid Energy System. *Sustainability*, 14(9), 5673-5702, 2022

<https://www.mdpi.com/2071-1050/14/9/5673>

(K.A. CARE project) Scopus indexed; Journal Rank: Q2; Impact factor: 3.889

12. **M. Maaruf**, and M. Khalid. Global Sliding-Mode Control With Fractional-

Order Terms for Robust Optimal Operation of the Hybrid Renewable Microgrid With Battery Energy Storage. *Electronics*, 11(1), 88-114, 2021

<https://www.mdpi.com/2079-9292/11/1/88/pdf>

Scopus indexed; Journal Rank: Q2; Impact factor: 2.690 (K.A. CARE project)

13. **M. Maaruf**, M.D. Shafiullah, Ali T. Al-Awami, and FSM Al-Ismail. Adaptive Nonsingular Fast Terminal Sliding Mode Control for Maximum Power Point Tracking of a WECS-PMSG. *Sustainability*, 13(23), 13427-13446, 2021

<https://www.mdpi.com/2071-1050/13/23/13427/pdf>

(Funded by Deanship of Scientific Research)

Scopus indexed; Journal Rank: Q2; Impact factor: 3.889

14. M.S. Mahmoud, **M. Maaruf**. Prescribed performance output feedback synchronisation control of bilateral teleoperation system with actuator nonlinearities. *International Journal of Systems Science*. 52(15), 3115-3127, 2021

<https://doi.org/10.1080/00207721.2021.1921308>

(Funded by Deanship of Scientific Research)

Scopus indexed; Journal Rank: Q2; Impact factor: 2.281; Acceptance rate: 32%

15. M.S. Mahmoud, **M. Maaruf**. Discrete-time attitude stabilization of reusable reentry vehicle by convex optimization. *Int. J. Dynam. Control*, 9(3),

1092-1099, 2021

<https://doi.org/10.1007/s40435-020-00738-4>

(My "Optimal and robust control" term project)

Scopus indexed; Journal Rank: Q2; Impact factor: 2.016

16. M.S. Mahmoud, **M. Maaruf**, & S. El-Ferik. Neuro-adaptive fast terminal sliding mode control of the continuous polymerization reactor in the presence of unknown disturbances. *Int. J. Dynam. Control*, 9(3), 1167-1176, 2021
<https://doi.org/10.1007/s40435-020-00731-x>

(From my master's thesis)

Scopus indexed; Journal Rank: Q2; Impact factor: 2.016

17. M.S. Mahmoud, **M. Maaruf**, S. El-Ferik. Neuro-adaptive output feedback control of the continuous polymerization reactor subjected to parametric uncertainties and external disturbances. *ISA transactions*, 11(12), 1-111, 2020
[https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S0019057820305139)

S0019057820305139

(From my master's thesis)

Scopus indexed; Journal Rank: Q1; Impact factor: 5.911

18. M. S. Mahmoud and **M. Maaruf**. Robust Adaptive Multilevel Control of a Quadrotor. In *IEEE Access*, 8, 167684-167692, 2020. <https://doi.org/10.1109/ACCESS.2020.3000000>

[//ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9199858](http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9199858).

(Funded by Deanship of Scientific Research)

Scopus indexed; Journal Rank: Q1; Impact factor: 3.367

CONFERENCE PAPERS

1. **M. Maaruf**, El-Ferik S, Al-Sunni FM. Neural Network-Based Control for Affine Formation Maneuver of Multi-Agent Systems with External Disturbances. In 2023 31st Mediterranean Conference on Control and Automation (MED) 2023 Jun 26 pp. 862-867. IEEE
(From my Ph.D. dissertation)
2. M. M. Gulzar, **M. Maaruf**, K. Khan, F. Alismail and M. Khalid, "Biologically Inspired Algorithm for Adaptive Convergence in Multi-Agent System," 2023 International Conference on Control, Automation and Diagnosis (ICCAD), Rome, Italy, 2023, pp. 01-05. IEEE
3. **M. Maaruf**, G. Konstantinou and M. Khalid, "Current Sharing Control Method for Fuel Cell Hybrid Electric Vehicles," 2022 IEEE PES 14th Asia-Pacific Power and Energy Engineering Conference (APPEEC), Melbourne, Australia, 2022, pp. 1-5. IEEE
4. **M. Maaruf**, S. El-Ferik, M. Abido and M. Khalid, "Two-Loop Distributed Optimal Secondary Voltage Control of an Islanded Microgrid with Multiple

- Distributed Generators,” 2022 IEEE PES 14th Asia-Pacific Power and Energy Engineering Conference (APPEEC), Melbourne, Australia, 2022, pp. 1-5. IEEE
5. **M. Maaruf**, S. El-Ferik and M. Khalid, ”A Decentralized Multiagent-Based Robust Backstepping Control for Restoring Secondary Voltage and Frequency of Autonomous Microgrids,” In 2022 11th International Conference on Renewable Energy Research and Application (ICRERA), 2022, pp. 430-435. IEEE
369-373
 6. **M. Maaruf**, S. El-Ferik, F. Saleh Al-Ismail and M. Khalid, ”Robust Optimal Virtual Inertia Control for Microgrid Frequency Regulation Considering High Renewable Energy Penetration,” 2022 11th International Conference on Renewable Energy Research and Application (ICRERA), Istanbul, Turkey, 2022, pp. 369-373. IEEE
 7. **M. Maaruf**, S. El-Freik, and Abdul-Wahid A. Saif, ”Hovering Control of a Quadrotor,” In 2022 19th International Multi-Conference on Systems, Signals & Devices (SSD),2022, pp. 2180-2185. IEEE
 8. **M. Maaruf**, M. Maaruf, S. Ahmad and M. Khalid, ”Improving the Frequency Response of High Renewable Penetrated Islanded Microgrids,” 2022 IEEE PES 14th Asia-Pacific Power and Energy Engineering Conference (APPEEC), Melbourne, Australia, 2022, pp. 1-6. IEEE

9. **M. Maaruf** and M. Khalid, "Hybrid Solar/ PEM Fuel Cell/ and Water Electrolyzer Energy System for All-Electric Ship," 2022 IEEE Kansas Power and Energy Conference (KPEC), 2022, pp. 1-5,
doi: 10.1109/KPEC54747.2022.9814802
10. **M. Maaruf** and M. Khalid, "Power Quality Control of Hybrid Wind/Electrolyzer/Fuel-Cell/BESS Microgrid," 2021 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia), 2021, pp. 1-5,
doi:10.1109/ISGTAsia49270.2021.9715639
11. **M. Maaruf**, K. A. Khan and M. Khalid, "Integrated Power Management and Nonlinear-Control for Hybrid Renewable Microgrid," 2021 IEEE Green Technologies Conference (GreenTech), 2021, pp. 176-180,
doi:10.1109/GreenTech48523.2021.00037
12. **M. Maaruf**, and S. E. Ferik, "Distributed Control Method for Heterogeneous Multiagent Systems Subjected to Faults," 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 2021, pp. 1328-1333,
doi:10.1109/SSD52085.2021.9429392. (My "Adaptive control" term project)
13. **M. Maaruf**, S. E. Ferik and M. S. Mahmoud, "Integral Sliding Mode Control

With Power Exponential Reaching Law for DFIG,” 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), 2020, pp. 1122-1127, doi:10.1109/SSD49366.2020.9364131. (My ”Control of nonlinear systems” term project)

14. Sami El-Freik, M. S. Mahmoud and **M. Maaruf**, ”Robust Adaptive Sliding Mode Control of Nonlinear Systems Using Neural Network,” 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), 2020, pp. 591-596, doi:10.1109/SSD49366.2020.9364137. (My ”Control of nonlinear systems” term project)

BOOK CHAPTERS

1. **Maaruf, Muhammad**: ’Neuro control of grid-connected three-phase inverters’ (Energy Engineering, 2023), ’Intelligent Control of Medium and High Power Converters’, Chap. 8, pp. 159-171,
DOI: 10.1049/PBPO239E_ch8 IET Digital Library,
https://digital-library.theiet.org/content/books/10.1049/pbpo239e_ch8