

Distributed Formation Control for Multi-agent Systems and Its Applications

by

Bing Yan

Thesis submitted for the degree of

Doctor of Philosophy

in

School of Electrical and Electronic Engineering
Faculty of Sciences, Engineering and Technology
The University of Adelaide

2022

© 2022
Bing Yan
All Rights Reserved



Supervisors:

Prof. Peng Shi, School of Electrical and Electronic Engineering

Prof. Cheng-Chew Lim, School of Electrical and Electronic Engineering

Contents

Contents	v
Abstract	vii
Originality Declaration	ix
Acknowledgments	xi
Publications	xiii
Chapter 1. Introduction	1
1.1 Background	2
1.1.1 Multi-agent systems	2
1.1.2 Formation control problems	3
1.1.3 Applications of multi-agent formations	4
1.2 Motivation	5
1.3 Summary of original contributions	7
1.4 Thesis structure	9
Chapter 2. Literature Review	11
2.1 Introduction	13
2.2 Publication	13
2.3 Supplementary literature review	29
2.3.1 Learning-based formation control	29
2.3.2 Collision-free formation optimization	30
Chapter 3. Robust Formation Control for Multi-agent Systems Based on Adaptive Observers	31
3.1 Introduction	34
3.2 Publication	34

Chapter 4. Robust Formation Control for Nonlinear Heterogeneous Multi-agent Systems Based on Adaptive Event-triggered Strategy	47
4.1 Introduction	49
4.2 Publication	49
Chapter 5. Optimal Robust Formation Control for Heterogeneous Multi-agent Systems Based on Reinforcement Learning	63
5.1 Introduction	66
5.2 Publication	66
Chapter 6. Event and Learning-based Resilient Formation Control for Multi-agent Systems under DoS Attacks	89
6.1 Introduction	92
6.2 Publication	92
Chapter 7. Conclusion	117
7.1 Summary	118
7.2 Future works	119
Bibliography	121

Abstract

MULTI-agent systems (MAS) consist of interacting entities, which can work together to solve complex problems that are difficult for an individual agent to possibly achieve. Formation control is a way to achieve collaborations in MAS by changing the motions of each agent and the distribution of the relative positions between agents.

In recent years, formation control for MAS, especially for heterogeneous MAS with different entities, has been intensively studied due to its wide range of applications in aerospace, intelligent transportation, and smart logistics. However, truly distributed, and reliable operations of MAS formations are difficult in practice with multiple constraints from interaction and physical systems. For instance, their interactive information is commonly locally incomplete and unreliable due to limited communication capabilities and potential cyber-attacks, and their physical systems are inevitably subject to unmodeled dynamics, dynamic barriers, etc. Therefore, the distributed and robust formation control for MAS is significant, and the transformation from control theoretical discoveries to real-world applications is essential.

In this thesis, a series of distributed formation control strategies are developed for heterogeneous multi-agent systems to ensure reliable operations, optimised performance, and flexible collaborations under interaction and physical system constraints. To evaluate the impacts of new strategies in practical systems, these discoveries are applied to autonomous vehicles in time-varying formations for target tracking and patrolling, collaborative collision avoidance, and area scanning.

First, formation control problems and methods for MAS are reviewed under two-layer constraints: 1) interaction layer constraints include local information, switching topologies, limited communication resources, cyber-attacks, etc. 2) physical system layer constraints include complex heterogeneous dynamics, multiple disturbances, uncertain even unknown model, limited real-time optimization and computing capabilities, physical barriers, etc. Then, we propose novel distributed adaptive observers, event-triggered mechanisms, and resilient control methods to guarantee the stability and resilience of MAS at the interaction layer. For physical system layer constraints, robust heterogeneous formation control, optimal collision avoidance algorithm, and

reinforcement learning-based model-free control strategies are provided to ensure safe operations, optimized performance, and flexible collaborations among different agents in dynamic environments. To address the practical collaborative problems, the developed control methods are applied in autonomous vehicles to perform collaborative tasks by dynamic formations. The results demonstrate the effectiveness, robustness, and resilience of the proposed strategies.

Originality Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

03 October 2022

Signed

Date

Acknowledgments

I would like to express my deepest gratitude to my principal supervisor Prof. Peng Shi for all his huge support, patient guidance, valuable suggestions, and constant encouragement in my professional development and personal growth. I am very grateful to him for giving me the opportunity and accepting me as a Ph.D student when I was at the most confusing time. During my three-year Ph.D study, He always patiently guides and encourages me to explore more in the research. His profound and comprehensive knowledge of control theories, autonomous systems, and artificial intelligence always inspire me. Extremely high standards and rigorous attitudes toward research are his iconic images. He always has a discussion with his Ph.D students about the research progress and new ideas to solve research problems. He is always happy and open to any discussions regarding research and life. It is his tremendous support and contributions that drive and push my Ph.D study to the finish line. He is not only a supervisor and mentor but also someone I will always look up to.

I also owe my deep gratitude to my co-supervisor, Prof. Cheng-Chew Lim, for his guidance, encouragement, and valuable feedback on the manuscripts. He always provides insightful discussions about machine learning and robotics. He always helps me polish my research writing skills. His response to emails and feedback on the academic manuscripts to his students are very timely even during nights and weekends. His rigorous research attitude always inspires me. I am appreciative of his invaluable guidance.

I want to acknowledge Yuan Sun for his help with research and experiments and Xin Yuan for laboratory facility support. In addition, I also owe my sincere gratitude to my friends and colleagues, Yutong Liu, Zhi Lian, Li Zhao, Yang Fei, Daotong Zhang, Hua Ma, Xiao Liu, Yuan Yuan and Xiaoyang Yin for their care, help and encouragement.

I would like to acknowledge the University of Adelaide, School of Electrical & Electronic Engineering, and the Australian Research Council for the scholarship and financial support.

Lastly, I would like to thank my family for their love and support. It is always encouraging knowing they are always there for me.

Publications

The following is a list of publications that are associated with the contents of this thesis. Among them, one publication is currently under review at the time of writing.

Journal Articles

- [1]. **B. Yan**, P. Shi and C. -C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2021.3103877, 2021.
- [2]. P. Shi and **B. Yan**, "A survey on intelligent control for multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 161–175, 2021.
- [3]. **B. Yan**, P. Shi, C. -C. Lim and C. Wu, "Robust formation control for multiagent systems based on adaptive observers," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3139–3150, 2022.
- [4]. **B. Yan**, P. Shi, C. -C. Lim and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 5, pp. 2683–2704, 2022.
- [5]. Y. Sun, **B. Yan**, P. Shi and C. -C. Lim, "Consensus for multi-agent systems under output constraints and unknown control directions," *IEEE Systems Journal*, accepted July 2022, doi: 10.1109/JSYST.2022.3192573.
- [6]. **B. Yan**, Y. Sun, P. Shi and C. -C. Lim, "Event and learning-based resilient formation control for multi-agent systems under DoS attacks," submitted and under review.

Conference Articles

- [1]. **B. Yan** and P. Shi, "Resource-aware formation control of uncertain multi-agent systems," in *Proceeding of Fourth International symposium on information and knowledge management (ISIKM2020)*, Online, December 12-13, 2020. **(Best Paper Award)**
- [2]. **B. Yan**, P. Shi, Z. Shi and H. Zhang "Learning-based collision-free formation control for heterogeneous multi-agent systems," in *Proceeding of 15th International Conference*

Publications

on Innovative Computing, Information and Control (ICICIC2021), Matsue, Japan, September 15-16, 2021. **(Best Innovation Paper Award)**

Chapter 1

Introduction

THIS chapter presents a brief background of multi-agent systems, formation control problems and applications of multi-agent formations. It is followed by the research motivation of this work and the original contributions of the thesis. At the end, the structural organization of the thesis is outlined.

1.1 Background

1.1.1 Multi-agent systems

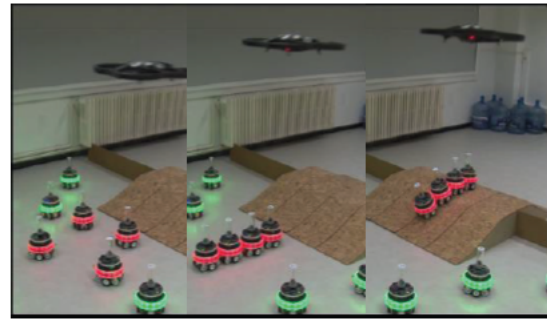
The concept of an agent emerged after the rise of artificial intelligence [1]. It refers to an entity operating in an environment, which is capable of autonomous action in the environment in order to meet its design goals [2,3]. The agents can be of any type, such as software [1], vehicles [4], drones [5], robotic arms [6], etc. There are three points worth noting from the concept: (1) autonomy means that an agent controls its own actions; (2) goal-directed behavior emphasizes that an agent changes its behavior to achieve its goals; and (3) the environment can be physical environment (e.g. in control engineering field) or computing environment (e.g. in computer science field). This thesis focuses on agents in control engineering.

Multi-agent systems (MAS) are systems composed of multiple interacting agents, which can work together to solve complex problems that are difficult for an individual agent to possibly achieve [2]. The advantage of MAS is delivered through (1) interactions between agents, and (2) collaborations in their actions to achieve some common goals. In control engineering, MAS can be considered as involving interaction and physical system layers (or levels) [7]. At the interaction layer, information is exchanged between individual agents through communication networks or/and sensor perception. At the physical system layer, each agent is a control system with physical characteristics, and it has its own influence areas in the shared environment [7].

From the composition of the physical system layer, MAS can be divided into homogeneous MAS [8–11] and heterogeneous MAS [12–16]. If the system is composed of identical agents, it is called homogeneous MAS. On the contrary, different agents with different dynamics consist of heterogeneous MAS. For example, a homogeneous MAS consisting of a group of Coachbots (V2.0) is developed to perform dynamic collaborative tasks by a task swapping algorithm, as shown in Fig. 1.1 (a) [8]. A heterogeneous MAS demonstrated in Fig. 1.1 (b) is composed of drones and ground-based self-assembling robots [12]. Drones use their privileged view of the environment to determine and communicate information to groups of robots on what morphologies to form to carry out upcoming tasks. Compared with homogeneous MAS, heterogeneous MAS show greater flexibility in collaborations because different agents have different capabilities to perform a collaborative task with less cost. From the engineering aspect, sometimes it is too difficult to equip the same agent with all the necessary modules.



(a) A homogeneous MAS consisting of Coachbots at Northwestern University, United States



(b) A heterogeneous MAS consisting of drones and robots at Université Libre de Bruxelles, France

Figure 1.1. Examples of MAS. (a) A homogeneous MAS consisting of Coachbots at Northwestern University, United States [8] (b) A heterogeneous MAS consisting of drones and robots at Université Libre de Bruxelles, France [12].

However, heterogeneous agents inevitably increase the complexity of multi-agent control system design.

1.1.2 Formation control problems

Formation control is a way to achieve collaborations in MAS by driving agents to maintain and move as desired geometric shapes without collisions [17–20]. Inspired by biological behaviors such as bird migration, MAS in formations increase the task execution efficiency, the adaptability to dynamic environments, and the survivability of the entire system.

The formation control problems involve how to control agents to reliably form geometric shapes, and how to ensure system safety without collisions. According to whether the desired geometric shape changes with time, the formation control problems can be divided into the fixed formation control problem [21, 22] and the time-varying formation (TVF) control problem [23–26]. TVF control for MAS improves the flexibility of their collaborations for executing dynamic tasks. From the structure of controller design, formation control problems can be classified as centralized and distributed formation control problems. As we can see from Fig. 1.2 (a), there is a control center in the centralized structure, which uniformly coordinates the actions of agents according to the collected global information and the centralized calculation [27]. As shown in

1.1.3 Applications of multi-agent formations

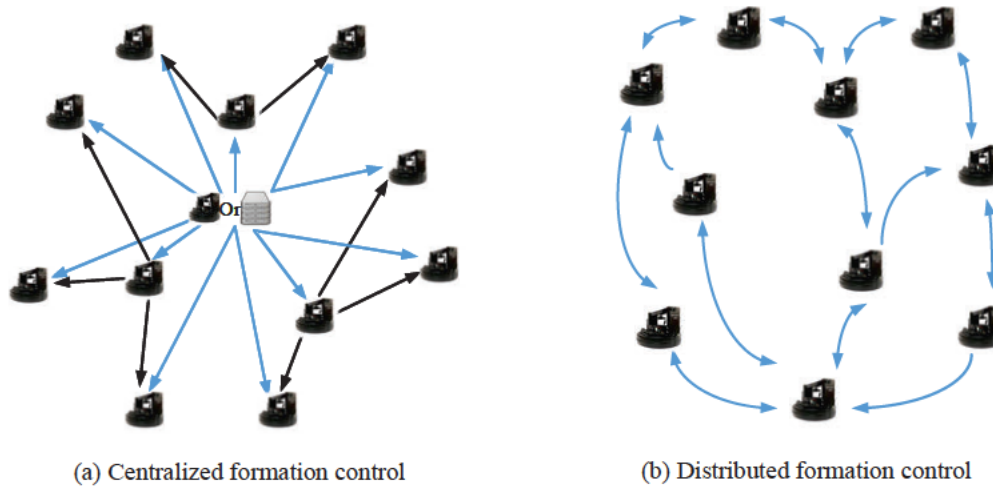


Figure 1.2. Formation control structures. (a) Centralized formation control (b) Distributed formation control.

Fig. 1.2 (b), the distributed formation control problem is to design a distributed control strategy for each agent to achieve some common formations only based on local information from their neighbors [11, 28, 29]. Distributed formation control enhances the resilience and robustness of MAS as the number of agents grows. This thesis focuses on distributed formation control problems for MAS, especially for heterogeneous MAS.

1.1.3 Applications of multi-agent formations

Multi-agent formations have a wide range of potential applications in aerospace [30, 31], robotics [32], intelligent transportation [33] and smart logistics [34]. For example, satellite formation flying based on System F6 in Fig. 1.3 (a) [30] aims to expand the overall coverage area and increase real-time performance, which can be used for positioning and navigation, weather monitoring, terrain exploration, etc. As shown in Fig. 1.3 (b), connected ground vehicles save energy in formations [34], which can be applied in smart logistics and intelligent transportation systems. Formation control is also being applied to unmanned vehicles, such as unmanned aerial vehicles (UAVs) [35], unmanned ground vehicles (UGVs) [36] and autonomous underwater vehicles (AUVs) [37] for collaborative area search, scanning, target tracking, monitoring, patrolling, 3-dimensional (3D) light display, disaster relief, search and rescue, etc. It reduces human investment and risks in some severe environments.

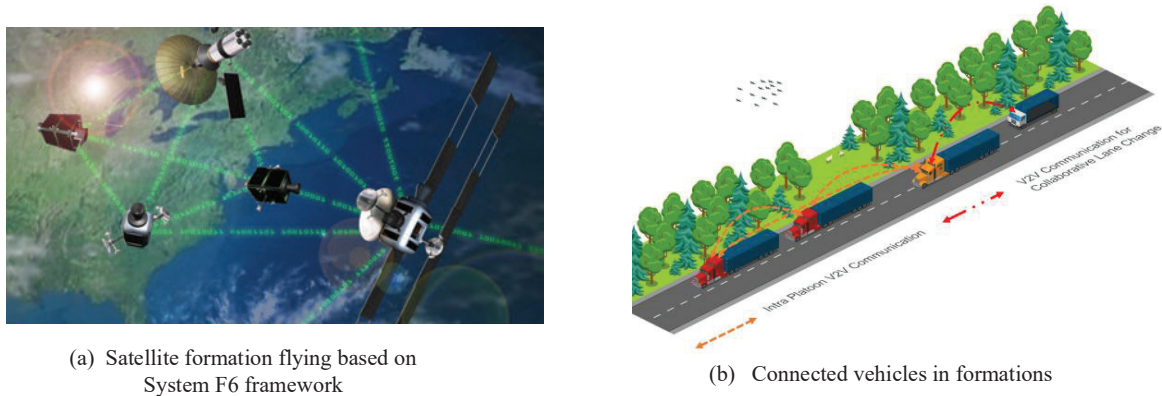


Figure 1.3. Multi-agent formation applications. (a) Satellite formation flying based on System F6 framework [30] (b) Connected vehicles in formations [34].

1.2 Motivation

As an attractive research area in control engineering and artificial intelligence, distributed formation technology contributes to efficient, flexible, fast, and powerful collaborations of MAS. However, fully distributed, and reliable operations of multi-agent formations are difficult in practice with multiple constraints from interaction and physical systems.

As shown in Fig. 1.4, the interaction layer constraints mainly include (1) incomplete local information as the communication capabilities of agents and the number of equipped devices are limited [29,38], (2) dynamic switching topologies on account of their bounded interaction range [39], (3) limited network resources and communication bandwidth [40], (4) unreliable cyber-environments due to potential cyber-attacks [41], and other network-induced issues.

Furthermore, there are inevitable constraints from the physical system layer of MAS, such as (1) complex heterogeneous dynamics because the agents in MAS may have different motion modes and state orders, (2) multiple disturbances from uncertain environments, (3) uncertain even unknown model information in practical systems, (4) limited real-time optimization and computing capabilities, (5) dynamic barriers in threat environments, etc.

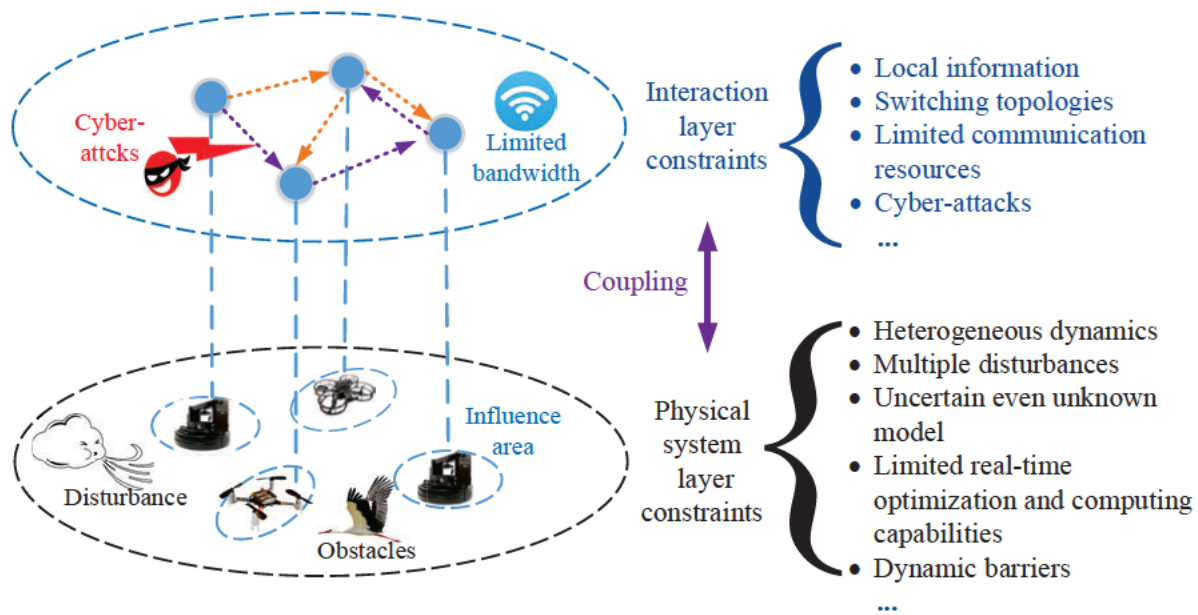


Figure 1.4. Two-layer constraints.

Multiple two-layer constraints threaten the stability, resilience, and robustness of MAS. Additionally, the distributed controller design involves coupling between the interaction layer and the physical system layer, because the controller of each agent requires the dynamics of neighboring agents, and the action of each agent under the controller affects each other in a shared environment. These constraints make the coupling problem more complicated.

Therefore, the research on distributed formation control for MAS under multiple constraints is significant. The transformation from control theoretical discoveries to real-world applications is essential. The distributed robust and resilient control technology is also of great significance to the research of cyber-physical systems [42] and human-machine collaborative control [43]. Together with communication and artificial intelligence technologies, it will promote the development of Industry 5.0 [44].

1.3 Summary of original contributions

This work aims to design distributed formation control strategies to ensure stability, robustness, and resilience of MAS under multiple constraints from interaction and physical system layers. Applying the control strategies to applications such as multi-UGV and UAV-UGV MAS is also our objective.

The original contributions of this work can be summarized as

1. A novel distributed robust control strategy is proposed for uncertain heterogeneous MAS to achieve TVF under switching topologies and multiple disturbances. Compared with existing methods for mixed-order heterogeneous MAS (e.g. a MAS composed of first-order integrators and second-order integrators), we propose a TVF control strategy for a unified linear heterogeneous MAS with different orders and dynamics to adapt to complex tasks. An adaptive observer is developed under switching topologies to estimate the state information of a reference exosystem only based on local information, which is used for decoupling the heterogeneous dynamics from networks. Considering the physical system layer constraints of uncertainties, homogeneous disturbances, and heterogeneous disturbances, a robust L_2 controller is designed for unified heterogeneous MAS to achieve TVF. A case study of a UAV-UGV TVF for bushfire edge tracking and patrolling is presented. Comparative simulation results demonstrate that our solution has significant advantages in the case of the MAS against multiple disturbances.
2. A brand-new dual adaptive TVF control scheme is proposed for nonlinear heterogeneous MAS to deal with limited network bandwidth constraints. Compared with linear MAS, a more general system, unified nonlinear heterogeneous MAS, is considered subject to uncertainties and disturbances. To reduce the frequency of data transmission, a distributed dual adaptive event-triggered observer is presented for exosystem estimation, which removes the global communication information in both observer design and Zeno-free event-triggered strategy design while saving network resources. A nonlinear p-copy internal model-based formation controller is designed with a dynamic distributed compensator for uncertainties and disturbances, which solves the robust nonlinear heterogeneous TVF problem. The scheme has been applied to a UAV-UGV MAS and a multi-UGV MAS for simulation and experimental verification. The results verify that the

1.3 Summary of original contributions

proposed scheme can significantly reduce communication frequency under the premise of ensuring the robustness of multi-agent formations.

3. Considering unknown heterogeneous MAS with an unknown exosystem, a novel reinforcement learning (RL)-based distributed formation optimization is provided to achieve TVF without collisions. Three new off-policy RL algorithms are proposed to learn the optimal policies of each agent in real time. An observed model-based RL algorithm or a model-free RL algorithm can be used to estimate the dynamics and states of a reference exosystem. Another model-free RL algorithm is integrated with a collision-free formation controller to solve TVF optimization problems in dynamic environments. Compared with most existing studies focusing on quadratic objective functions, the developed control method addresses the non-quadratic optimization problem when the system model is completely unknown. Comparative simulations demonstrate the real-time learning performance and dynamic collision avoidance capability of a UAV-UGV heterogeneous MAS.
4. Considering one of the typical cyber-attacks, the denial-of-service (DoS) attack, we propose a novel resilient and robust two-layer controller design with a brand-new RL condition to address TVF problems for unknown heterogeneous MAS. The design is distributed and model-free at the cyber-layer and the physical system layer. We specify the interaction layer as the cyber-layer as we focus on network interaction rather than sensing interaction. An event-based resilient observer is provided at the cyber-layer to remove global information of communication and deal with DoS attacks. The communication load can be reduced under attacks, and the Zeno behavior can be avoided. In the physical system layer, an RL rank condition for the TVF controller is developed for unknown heterogeneous MAS. Compared with most existing works on off-policy RL for heterogeneous MAS, the new rank condition can automatically adjust online data collection time, thereby improving online learning and optimization performance. Experiments of multi-UGV area scanning formations are conducted. The comparative experimental results verify the resilience of the proposed online event and learning-based control method under different parameters of DoS attacks.

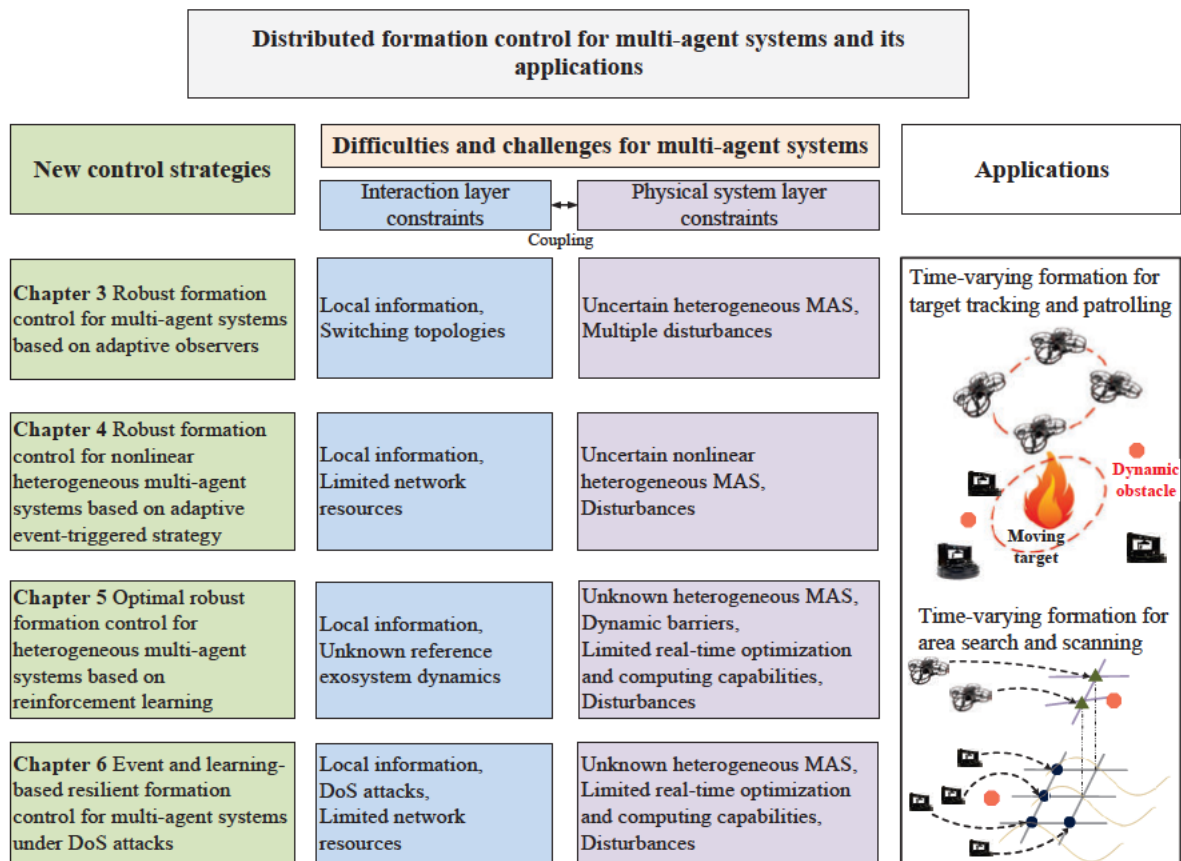


Figure 1.5. Thesis outline.

1.4 Thesis structure

This thesis starts with introduction and literature review in the first two chapters and ends up with a conclusive and foreseeing discussion in the last chapter. As illustrated in Fig. 1.4, the main chapters in between focus on the original contributions of the thesis, as described in detail in the following.

- Chapter 1 introduces the background of MAS, formation control problems and applications of multi-agent formations followed by the motivation, the contribution summary and the outlines of the thesis.
- Chapter 2 provides a survey on intelligent control of MAS, where the distributed formation control methods under two-layer constraints have been reviewed. Then, a supplementary literature review on formation control with some new methods since the published survey release is added.

1.4 Thesis structure

- In Chapter 3, an adaptive observer-based robust L_2 formation control strategy is proposed for uncertain heterogeneous MAS under multiple disturbances and switching topologies. It has been applied to a UAV-UGV MAS to execute target tracking and patrolling tasks with simulation verification.
- Chapter 4 provides a dual adaptive event-trigger-based robust formation control for uncertain nonlinear heterogeneous MAS under limited network resources. Both simulations and experiments are given for multi-vehicle TVF to verify the effectiveness of the proposed strategy.
- In Chapter 5, an optimal model-free and collision-free formation control scheme is developed for unknown heterogeneous MAS with unknown exosystem dynamics based on RL. It has been applied to a UAV-UGV MAS for simulation verification.
- Chapter 6 presents an event and learning-based resilient formation control for unknown heterogeneous MAS under DoS attacks. We apply the control method to a multi-UGV system for area scanning TVF with experimental verification.
- Chapter 7 provides a conclusive summary of the thesis, as well as a prospective view of the further work.

This thesis presented our original findings in a thesis by publication format. As such, the next five chapters are the five scholarly publications that resulted from my Ph.D.

Chapter 2

Literature Review

2.1 Introduction

This chapter first presents a survey on distributed intelligent control of multi-agent systems (MAS) from the perspective of different constraints at interaction and physical system layers (or levels). Multiple constraints from two layers such as incomplete local information, system uncertainties, and limited interaction capabilities affect the performance of the entire MAS. A review is conducted on the development of MAS intended for intelligent control, including consensus problem, formation control, and flocking control. Based on the two-layer constraints, the research results on intelligent control are categorized into limited sensing-based control, event-based control, pinning-based control, resilient control, collaborative control for homogeneous MAS and collaborative control for heterogeneous MAS. The applications of intelligent control for MAS are reviewed and a discussion about the challenges is presented.

Then, a supplementary literature review on formation control with some new methods since the publication release is provided with evaluations.

2.2 Publication

P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 161–175, 2021.

A Survey on Intelligent Control for Multiagent Systems

Peng Shi¹, Fellow, IEEE, and Bing Yan

Abstract—In practice, the dual constraints of limited interaction capabilities and system uncertainties make it difficult for large-scale multiagent systems (MASs) to achieve intelligent collaboration with incomplete local relative information. In this article, a review is conducted on the recent development of MASs intended for intelligent control, including consensus problem, formation control, and flocking control. Based on the limitations of the interaction level and the constraints of the individual system level, the published results on intelligent control are categorized into limited sensing-based control, event-based control, pinning-based control, resilient control, and collaborative control under system constraints. Also, the applications of intelligent control for MASs are presented, especially for robotics, complex networks, and transportation. Finally, a discussion is given about the challenges and future directions of research in this field.

Index Terms—Consensus problem, flocking control, formation control, intelligent control, multiagent systems (MASs), multilevel constraints.

I. INTRODUCTION

MULTIAGENT collaborative intelligence technology has led to revolutionary changes for the practical applications of robotics, complex networks, and transportation in recent years [1]–[3]. In the meantime, however, it also presents challenges as distributed and reliable intelligence technology is required to perform cooperation tasks for large-scale multiagent systems (MASs) in the context of incomplete local relative information. Inspired by the collaborative behaviors observed in nature, such as bird migration in groups and flocking behaviors of fish schools, collaborative awareness, task assignment, and intelligent control of MASs have attracted much attention from various fields over the past decades [4]–[6]. As the fundamental way to ensure the successful collaborative missions for MASs, advanced intelligent control strategies are the focus in this survey article.

From the perspective of computer science, an agent refers to a computing system operating in an environment with certain levels of autonomy and capability of sensing, decision making, and acting [1]. As shown in Fig. 1, there are four

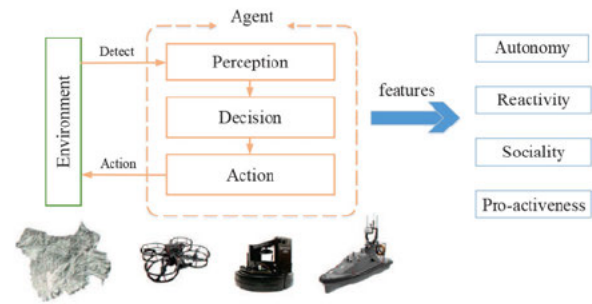


Fig. 1. Agent and its features.

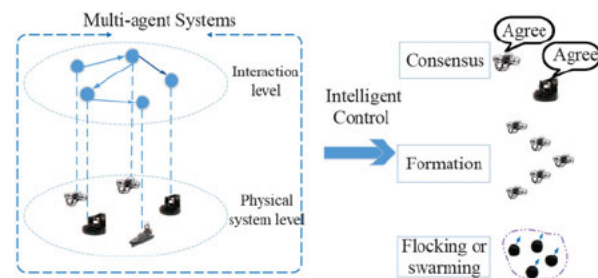


Fig. 2. Intelligent control for MASs.

features commonly used to describe an agent, including autonomy, reactivity, sociality, and proactiveness. Autonomy means that an agent is capable to operate without the direct intervention of other entities and exercising control over its own actions. Reactivity is defined as the capability of an agent to make a response to the changes in the environment and convert its sensory inputs to actions. Sociality means that an agent has the capability to make communications with others, while proactiveness stands that an agent can do more than acting in response to the environment. However, the capability of a single agent is limited, especially in dealing with complex tasks.

A MAS refers to a group of agents. It is capable to interact, coordinate their behavior, and cooperate to achieve some common goals [1]. In comparison with single-agent systems, MASs provide a more effective and robust way to solve various complex problems by means of collaborative intelligence. As shown in Fig. 2, a MAS consists of an information interaction level and a physical system level. At the interaction level, information is exchanged not only between individual agents but also between agents and their ambient environment through either communication networks or sensor perception.

Manuscript received November 19, 2020; revised November 25, 2020; accepted December 1, 2020. Date of publication December 24, 2020; date of current version January 12, 2021. This work was supported by the Australian Research Council under Grant DP170102644. This article was recommended by Associate Editor C. K. Ahn. (Corresponding author: Peng Shi.)

The authors are with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia (e-mail: peng.shi@adelaide.edu.au; bing.yan@adelaide.edu.au).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TSMC.2020.3042823>.

Digital Object Identifier 10.1109/TSMC.2020.3042823

2168-2216 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

In practice, the range of communication capabilities and perception is limited for a single agent. At the physical system level, constraints, such as uncertainties and complex heterogeneous dynamics, can have a severe impact on the performance of MASs.

As for the work on intelligent control of MASs, it focuses on three major problems, including consensus problem, formation control problem, and flocking/swarming problem. As the foundation of research on collaborative intelligence, the consensus problem of MASs has now been extensively investigated for all agents to achieve a common goal [3], [7]–[11]. As an extension to consensus problems, formation control aims at driving intractable agents to maintain and move as desired geometric shapes to perform predefined tasks, such as effective search, patrol, and exploration [6], [12], [13]. As a self-organizing behavior, flocking or swarming is derived from small-size animals with lower intelligence [14]–[16], for example, bees, fish school, and bird swarms, in the process of migrating, cruising, or avoiding enemies. Swarming is also extended to describe the behavior of lifeless agents, such as robots. Swarm intelligence not only expands individual capability but also improves the overall level of survivability.

Depending on different control structures, the approaches taken for the intelligent control for MASs can be classified into centralized control [17] and distributed control [18]. In respect of centralized control, there is a control center or host in place to coordinate the information transmission and the final process of task completion. However, the failure of the control center will hinder the entire system from functioning as normal. To improve the robustness of the whole MASs, distributed control approaches have been widely studied, where all agents determine their behavior separately based on local information. The design of distributed intelligent control that only relies on incomplete local information has been a study focus in recent years [2], [7], [18].

Information interaction and system dynamics play a vital role in the intelligent control for the entire MASs. At the interaction level, the information interaction between agents relies on the capabilities of each agent to carry out sensing and communication. In general, the information level of practical MASs is subject to a limited range of perception without communication, limited bandwidth with communication, limited network resources, and other network-induced issues. According to the different limitations on the interaction level, the recently proposed methods to solve MAS collaborative control can be categorized into limited sensing-based control [19], event-driven control [20], pinning control [21], and resilient control [22].

At the system level, an overview of theoretic advancements in consensus, formation control, and flocking control has been presented for MASs with single-integrator and double-integrator dynamics under fixed and switching topologies in [11] and [23]–[25]. In addition to linear MASs, a large number of works on reliable intelligent control have been intensively investigated for homogeneous nonlinear MASs under systems uncertainties and disturbances [7], [26]–[28]. Compared with homogeneous agents, heterogeneous agents show greater flexibility in task allocation depending on

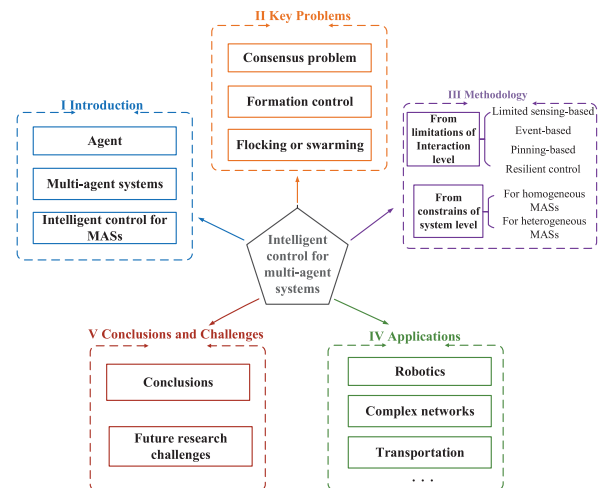


Fig. 3. Structure of the article.

different capabilities in cooperative operations. As one of the basic heterogeneous systems, hybrid-order MASs consist of different order integrator systems were described in [29]–[32]. However, the less restrictive heterogeneous systems with different orders and different dynamics are more commonly applied in practice. In recent years, the robust output regulation control exercised by general heterogeneous MASs has attracted a great number of attention [33]–[36]. Subsequently, it was extended to solve the formation control problem and the flocking problem encountered by heterogeneous systems. The challenge still arises from the systems due to the limitations of uncertainties and heterogeneous dynamics.

In this article, a survey is conducted on the recent study of MASs in intelligent control considering the constraints of information interaction level and system level. We try our best to summarize the relevant research work in recent years, and apologize for missing some contributions on the topic, if any.

The overall structure of the article is shown in Fig. 3. The background and basic concepts are introduced in Section I. Preliminaries and three major problems with intelligent control for MASs are described in Section II, including consensus, formation, and flocking problems. In Section III, the recent advancement of the methodologists to tackle the underlying problems is described with the constraints of the interaction level and system level. Then, a review is presented on the main applications of robotics, complex networks, and transportation in Section IV. The conclusion and challenges ahead are discussed in Section V.

II. PRELIMINARIES AND KEY PROBLEMS OF INTELLIGENT CONTROL FOR MASS

In this section, preliminaries about graph theory and three key issues of intelligent control are recalled, including consensus problem, formation control, and flocking or swarming.

In general, the information exchange among agents is modeled by a directed or undirected graph [37]. In a graph $G = (V, E)$, V represents a finite nonempty set of nodes and E is an edge set, which contains ordered pairs of nodes in

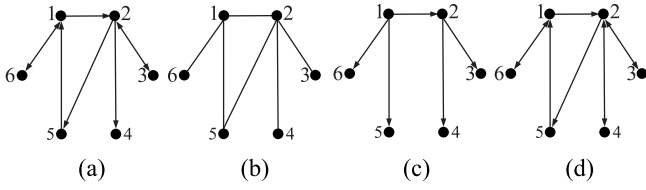


Fig. 4. Graphs with different features. (a) Directed graph. (b) Undirected graph. (c) Spanning tree. (d) Connected graph.

a directed graph and unordered pairs of nodes in an undirected graph. For example, $(i_1, i_2) \in E$ indicates that i_2 obtains the information from i_1 in a directed graph [Fig. 4(a)], and i_1 and i_2 can obtain the information from each other in an undirected graph [Fig. 4(b)]. The neighbor set of agent i is $N_i = \{j \in V | (j, i) \in E\}$. A graph contains a spanning tree if there is a path between one node and all other nodes [Fig. 4(c)]; a graph is connected if there is a path between every pair of distinct nodes [Fig. 4(d)]. For a graph G , an adjacency matrix $A = [a_{ij}]$ specifies the interconnection topology of MASs, where

$$a_{ij} = \begin{cases} 0 & i = j, \text{ or } (j, i) \notin E \\ 1 & (j, i) \in E. \end{cases} \quad (1)$$

The Laplacian matrix L of graph G is $L = D - A$ where $D = \text{diag}(d_1, d_2, \dots, d_n)$ is the degree matrix with diagonal elements $d_i = \sum_j a_{ij}$.

A. Consensus Problem

As one of the research foundations for intelligent control for MASs, consensus refers to all systems reaching an agreement on certain interests regarding to their states and the concept comes from distributed computing systems and management science. A typical consensus theoretical framework was presented in [11], and the framework established some communication rules among the agent and their neighbor agents in the networks in order to achieve a common goal. The study also emphasized that graph theory and Laplace matrix were the core means to solve the consensus problem. Another work linked the minimum spanning tree theory in the graph with the information consensus framework [23] and proposed the minimum necessary and sufficient condition of information consensus for MASs under changing topologies, which laid the foundation for the research of dynamic topologies.

According to the theoretical frameworks, consensus problems can be divided into leaderless consensus problem and leader-follower consensus problem.

Problem 1: A general leaderless consensus problem is to design a controller for a MAS to meet

$$\lim_{t \rightarrow \infty} \|z_i(t) - z_j(t)\| = \mathbf{0}, \quad j \in N_i \quad (2)$$

where $z_i(t) \in \mathbb{R}^m$ and $z_j(t) \in \mathbb{R}^m$ represent the state or output of i th agent and j th agent, respectively. N_i is the neighbor set of agent i .

Problem 2: A general leader-follower consensus problem is to design a controller for a MAS to meet

$$\lim_{t \rightarrow \infty} \|z_i(t) - z_0(t)\| = \mathbf{0}, \quad i = 1, 2, \dots, n \quad (3)$$

TABLE I
RECENT WORKS ON CONSENSUS PROBLEMS

Features	Classification	References
From consensus structure	Leaderless consensus	[38]–[40]
	Leader-follower consensus	[40]–[46]
From interaction level	Event-based	[9], [47]–[57]
	Pinning-based	[53], [58]–[60]
	Resilient control	[61], [62]
From system level	Homogeneous linear	[38], [39], [63]
	Homogeneous nonlinear	[7], [27], [40], [45], [64]
	Heterogeneous linear	[42], [65], [66]
	Heterogeneous nonlinear	[30], [67]–[69]

where $z_i(t) \in \mathbb{R}^m$ represents the state or output of agent i . $z_0(t) \in \mathbb{R}^m$ is a common desired trajectory for all agents to track asymptotically.

Remark 1: It is worth noting that the leader can be a real physical system, or a virtual reference system designed according to the tasks. The final consensus of all agents without reference to their initial conditions in the leader-follower consensus problem. In terms of the leaderless consensus problem, all agents finally reach a consensus, which is related to the initial state of the system and the information interaction topology.

A summary of recent works on Problems 1 and 2 is outlined in Table I, corresponding to different features and constraints from system level and interaction level.

B. Formation Control Problem

Formation control is designed to drive the moving interacting agents to achieve or maintain a specified geometry for a coordinated goal. The formation control problem can be uniformly summarized into a consensus-based structure [70], after considering the reference formation dynamics and the motion characteristics of the agent.

Most results on formation control have focused on two main problems: 1) leaderless formation control problem and 2) formation tracking problem.

Problem 3: A general leaderless formation control problem is to design a controller for a MAS to meet

$$\lim_{t \rightarrow \infty} \|(z_i(t) - z_j(t)) - (f_i - f_j)\| = \mathbf{0}, \quad j \in N_i \quad (4)$$

where $(f_i - f_j)$ represents the reference formation deviation between agent i and agent j .

Problem 4: A general formation tracking problem is to design a controller for a MAS to meet

$$\lim_{t \rightarrow \infty} \|z_i(t) - f_{i0} - z_0(t)\| = \mathbf{0}, \quad i = 1, 2, \dots, n \quad (5)$$

where $f_{i0} \in \mathbb{R}^m$ is a referent formation deviation regards to a desired trajectory $z_0(t) \in \mathbb{R}^m$. If f_{i0} is a dynamic formation variable, the problem is extended to a time-varying formation problem.

Remark 2: Note that formation control Problems 3 and 4 can be viewed as extensions of consensus Problems 1 and 2, respectively, with respect to a reference formation deviation. For example, as shown in Fig. 5, the agent i needs to maintain a diamond formation deviation $(f_i - f_j)$ with its neighbor agent j , in leaderless formation control Problem 3. Therefore, the main

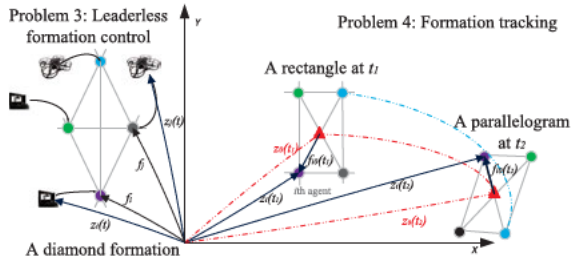


Fig. 5. Leaderless formation problem and formation tracking problem.

TABLE II
RECENT WORKS ON FORMATION CONTROL

Features	Classification	References
Formation problems	Leaderless formation	[19], [74]–[77]
	Formation tracking	[78]–[83]
From interaction level	Limited sensing-based	[19], [75], [84]–[86]
	Event-based	[13], [22], [74], [83], [87]
	Pinning-based	[87], [88]
	Resilient control	[74], [89], [90]
From system level	Homogeneous linear	[91], [92]
	Homogeneous nonlinear	[22], [74], [81], [88], [93]
	Heterogeneous linear	[32], [77], [82]
	Heterogeneous nonlinear	[79], [80], [94]–[96]

focus of Problem 3 is to form a formation. As for formation tracking Problem 4, there are a moving target $z_0(t)$ to track and a formation deviation f_{i0} to keep for agent i . In Fig. 5, a MAS aims to track a desired trajectory $z_0(t)$ from a rectangle to a parallelogram.

The approaches for formation control reported in the literature include the leader–follower control [12], [71], the behavior-based control [72], and the virtual structure approach [73]. In the leader–follower approach, the controller relies heavily on a single leader state. For the behavior-based formation method, several basic control behaviors of the agent are defined and weighted to obtain the final formation control inputs for the group. However, group behaviors are difficult to define. In the virtual structure approach, the formation of agents is regarded as a single object in the virtual structure, which limits the application domain as it only controls the motion of one object. The existing results can also be divided into position-based, displacement-based, and distance-based control, according to the sensing capability and the interaction topology of MASs [6]. Focusing on the constraints and features from the system level and the interaction level of agents, we list some related works on formation control in Table II.

C. Flocking or Swarming Control Problem

The flocking or swarming control problem of MASs is to perform a macroscopic overall synchronization, such as aggregate together and maintain the same direction, by using local interaction and behavioral rules between agents. The Boid model was first proposed based on computer simulation technology to describe bird swarms, which follows three rules: 1) cohesion: remain close to neighbors; 2) separation: avoid collision with neighbors; and 3) alignment: match velocity with neighbors [97]. The work in [98] simplified the Boid model and described the behavior of birds as a discrete model,

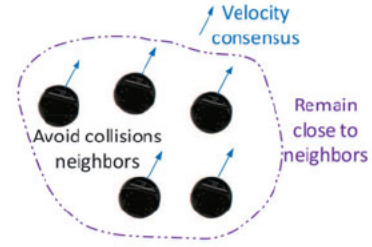


Fig. 6. Flocking or swarming control problem.

the Vicsek model. Its advantage is to quantitatively analyze the flocking behavior by changing the population density and noise intensity, from the point of view of statistical mechanics. These models provide a solid foundation for the research of the flocking problem.

According to the model and rules, the flocking control problem can be described as follows.

Problem 5: A general flocking/swarming control problem is to design a controller for a MAS to meet the three following rules.

- 1) *Cohesion:* Remain close to neighbors

$$\lim_{t \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n \|p_i(t) - p_j(t)\| < R, \quad j \in N_i \quad (6)$$

where $p_i(t) \in \mathbb{R}^m$ and $p_j(t) \in \mathbb{R}^m$ represent the positions of i th agent and j th agent, respectively. R is the maximum value of the sum of relative distances between agents.

- 2) *Separation:* Avoid collision with neighbors

$$\lim_{t \rightarrow \infty} \|p_i(t) - p_j(t)\| \geq d_{ij}, \quad j \in N_i \quad (7)$$

where d_{ij} is the minimum safety distance between agent i and agent j .

- 3) *Alignment:* Match velocity with neighbors

$$\lim_{t \rightarrow \infty} \|v_i(t) - v_j(t)\| = \mathbf{0}, \quad j \in N_i \quad (8)$$

where $v_i(t) \in \mathbb{R}^m$ and $v_j(t) \in \mathbb{R}^m$ denote the velocities of agent i and agent j , respectively.

To solve the problem according to the three rules shown in Fig. 6, a theoretical framework was proposed for the design of distributed flocking algorithms of second-order MASs [25]. The work in [99] analyzed the stability properties of flocking algorithms for second-order MASs under switching networks. We summarize some recent works on flocking and swarming problems in Table III, corresponding to different features and constraints from the system level and the interaction level.

III. METHODOLOGIES OF INTELLIGENT CONTROL FOR MASS

In this section, we review the results reported for the collaborative intelligence of MASs and outline advanced methodologies based on the limitations of information interaction level and the constraints of system level, respectively.

Passive sensing and active communication are the two important means of information interactions. For example, if

TABLE III
RECENT WORKS ON FLOCKING OR SWARMING CONTROL

Features	Classification	References
From interaction level	Limited sensing-based	[15], [16], [100]
	Event-based	[101]–[106]
	Pinning-based	[107], [108]
	Resilient control	[109]
From system level	Homogeneous linear	[110], [111]
	Homogeneous nonlinear	[101], [102], [112]–[117]
	Heterogeneous linear	[118]
	Heterogeneous nonlinear	[119], [120]

each agent is equipped with advanced sensors, which make them able to detect the location of neighbors, then communication is not required in formation control. However, the detection range of a single agent is usually limited. It becomes relatively easy for collaboration if agents can directly interact with key information through the network, such as locations and velocities. While communication networks are restricted by limited bandwidth, limited resources, and other network-induced issues. From the limitations of information interaction, we focus on reviewing the following popular and important topics: limited sensing-based control, event-based control, pinning-based control, and resilient control.

A. Limited Sensing-Based Control

Under certain situations such as an electrostatic shielding environment, the communication between agents is unavailable or very limited. Sensor-based perception is another alternative to achieve multiagent collaboration. In fact, it is beneficial to realize distributed control if agents can completely perceive neighbors and the environment independently. Moreover, the system can be immune to any network problems due to the independence of the communication network. However, for a single agent, the detection range of sensors, such as cameras and infrared sensors, is locally limited. Note that range-only sensing agents usually refer to robotics rather than network agents. Therefore, most of the existing results on limited sensing-based control have focused on solving formation control problems and flocking control problems [25], [75]. As shown in Fig. 7, formation reference denotes static or dynamic predefined displacement references regard to different formation shapes, while flocking reference refers to a moving rendezvous point as the group objective [25]. There are two key issues: 1) how does the agent determine the most suitable position relative to its neighbors? and 2) how to achieve collaboration actions without collisions?

In order to solve the two problems, a novel scalable formation control strategy was proposed to solve Problem 3, when the communication network is completely unreachable [75]. Consider a first-order MASs under a control input $u_i(t)$

$$\begin{aligned} \dot{p}_i(t) &= u_i(t), \quad i = 1, 2, \dots, n \\ u_i(t) &= f(N_i, g(N_i, F)) \end{aligned} \quad (9)$$

where $p_i \in \mathbb{R}^m$ and $u_i \in \mathbb{R}^m$ present the position and control input of the i th agent, respectively. Suppose the sensing range is limited, the neighbor set is defined as

$$N_i = \{ (j, i) \in E \mid \|p_i(t) - p_j(t)\| < r_i \}$$

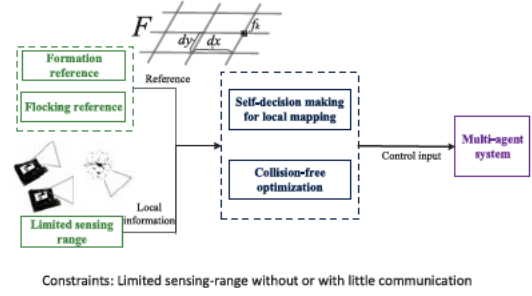


Fig. 7. Limited sensing-based intelligent control.

where r_i is the sensing range of the i th agent, which can be different for heterogeneous agents. The control force f aims to form a reference formation F , which is a set of scalable displacement vectors $F = \{\dots, f_{-1}, f_0, f_1, \dots\}$. The element $f_k \in \mathbb{R}^m$ is based on m linearly independent basis vectors, where $m = 2$ or $m = 3$ indicates the reference formation is in two dimensions or three dimensions. For instance, a 2-dimensional (2-D) scalable formation F is shown in Fig. 7, where d_x and d_y are two linearly independent vectors. For any $f_k \in F$, $f_k = ad_x + bd_y$, $a, b \in \mathbb{Z}$ always holds. Note that the basis vectors of F are known for all agents. The local mapping function g is conducted on each agent to choose which vectors in F are the suitable displacement according to the neighbor set N_i . Based on local optimization methods, multiobject mapping protocols without conflict and range control strategies have been investigated in [19], [75], and [84]. In addition to collision-free matching, obstacle avoidance algorithms have also been extended to the formation control in Problems 3 and 4 without communication [121] and with limited communication only on identities and mapping decisions [76]. In order to reduce the computation burden and avoid the infinite trajectory loop, the probability was introduced in mapping and distributed control strategies [85], [86].

As for the flocking control in Problem 5, most of the works have focused on the implementation of three rules for agents subject to a limited range-based perception or interaction. A theoretical flocking framework was proposed of a second-order MASs [25]

$$\dot{p}_i(t) = v_i(t), \quad \dot{v}_i(t) = u_i(t) \quad i = 1, 2, \dots, n \quad (10)$$

under the control force

$$\begin{aligned} u_i(t) &= f_i^g + f_i^d + f_i^f \\ &= \sum_{j \in N_i} \psi(\|p_j - p_i\|_\sigma)(\mathbf{n}_{ij}) + \sum_{j \in N_i} (v_j - v_i) \\ &\quad + c_1((p_0(t) - p_i(t))) + c_2((v_0(t) - v_i(t))) \end{aligned} \quad (11)$$

where the system state $x_i(t) = [p_i(t), v_i(t)]^T$ consists of position $p_i \in \mathbb{R}^m$ and velocity $v_i \in \mathbb{R}^m$ of the i th agent. To implement the three rules, control input consists of three items: 1) gradient-based term f_i^g ; 2) velocity consensus item f_i^d ; and 3) navigational feedback term f_i^f , where ψ is a potential function based on σ -norm of the relative distance for collision avoidance. (\mathbf{n}_{ij}) represents a vector connecting agent i and agent j . Parameters c_1 and c_2 denote the position control law

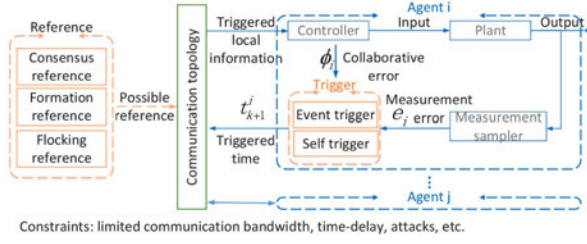


Fig. 8. Event-based intelligent control.

and velocity control law of the navigational feedback, respectively. They are designed to meet $c_1, c_2 > 0$. The flocking reference position $p_0 \in \mathbb{R}^m$ and velocity $v_0 \in \mathbb{R}^m$ indicate a rendezvous point, which can be viewed as a group objective. The connectivity preservation of formation control was also been discussed in [25].

Based on the basic flocking framework, flocking control solutions to Problem 5 were provided based on potential field [14] and learning-based approach [15], [16]. Without a communication channel, learning vision-based flocking algorithms was proposed for multidrone swarms [15]. Flocking control was developed for a first-order MAS subject to limited heterogeneous interaction range [100]. However, connectivity preservation is still a challenge for high-order nonlinear MASs, when the sensing capability of the agent is insufficient.

B. Event-Based Intelligent Control

The communication-based MASs make up for the lack of perception of some agents and effectively realize collaboration by directly interacting with the information of interest through the network. However, limited network resources and communication bandwidth have largely restricted information transmission between agents. Event-based distributed interaction mechanisms were proposed to alleviate this limitation [9], [47]–[57]. Without collecting the state information of all agents at every moment, the distributed event-triggered strategy can better save communication resources and effectively reduce the frequency of information transmission between agents and update of agent control protocols.

The structure of event-based distributed intelligent control is shown in Fig. 8, where each agent independently determines its own behaviors. The trigger determines the interaction time interval of each agent according to the measurement error from the sampler and the collaborative error from the controller. The controller updates the local information and possible reference information at each trigger moment. Note that the possible reference information here refers to the consensus reference, or the formation reference, or the flocking reference information. For the leaderless consensus problem in Problem 1, there is no external reference information. The event-based distributed trigger strategy mainly involves three key issues: 1) how to determine the trigger time; 2) how to design the distributed control laws; and 3) how to exclude the unlimited trigger phenomenon, Zeno phenomenon [122]. We take the event-based control for second-order MASs (10) as an example.

1) *Trigger Mechanism*: It is designed to determine the trigger time in next step t_{k+1}^i . The recent works mainly focus on two common mechanisms, event-triggered mechanism, and self-triggered mechanism, as follows:

$$\begin{aligned} (1) \quad t_{k+1}^i &= \inf\{t > t_k^i | f(e_i) > g(e_i, \phi_i)\} \\ (2) \quad t_{k+1}^i &= t_k^i + \sigma_i \end{aligned} \quad (12)$$

where $f(e_i)$ and $g(e_i, \phi_i)$ are the trigger functions based on measurement error e_i and collaboration error ϕ_i . The next time is triggered if the condition is met. Note that the trigger functions are only based on local information without the prior information of the topology matrix by adding adaptive laws in $f(e_i)$ or $g(e_i, \phi_i)$ [47]–[50]. The second mechanism is the self-triggered strategy, where the time interval of next broadcast σ_i can be calculated based on the information of the current trigger time without the need to continuously monitor the changes of events [104]. Although event monitoring costs are reduced, additional computational costs are added.

There are two common designs of measurement error

$$\begin{aligned} (1) \quad e_i(t) &= \hat{x}_i(t) - x_i(t) = x_i(t_k^i) - x_i(t) \\ (2) \quad e_i(t) &= \hat{x}_i(t) - x_i(t) = e^{A(t-t_k^i)} x_i(t_k^i) - x_i(t) \end{aligned} \quad (13)$$

where t_k^i is the k th trigger time of agent i , and $\hat{x}_i(t)$ represents the estimation of x_i in the time period $t \in [t_k^i, t_{k+1}^i)$. The state value at the time of the last trigger is used in the first strategy [48], [51]–[54]. The second approximation scheme of $x_i(t)$ is based on the system state matrix A . System model-based estimation more accurately approximates the state of the system during the time period [49], [50].

The design of the collaboration error is based on different intelligent control problems. For leader–follower consensus Problem 2, formation control in Problem 4, and flocking tracking control in Problem 5, the collaborative errors are generally formed as

$$(1) \quad \phi_i(t) = \sum_{j \in N_i} a_{ij}(\hat{x}_j(t) - \hat{x}_i(t)) + a_{i0}(\hat{x}_0(t) - \hat{x}_i(t)) \quad (14)$$

$$\begin{aligned} (2) \quad \phi_i(t) &= \sum_{j \in N_i} a_{ij}((\hat{x}_j(t) - \hat{x}_i(t)) - (\hat{f}_j(t) - \hat{f}_i(t))) \\ &+ a_{i0}((\hat{x}_0(t) - \hat{x}_i(t)) - (\hat{f}_{i0}(t) - \hat{x}_i(t))) \end{aligned} \quad (15)$$

$$\begin{aligned} (3) \quad \phi_i(t) &= \sum_{j \in N_i} a_{ij}(\hat{v}_j(t) - \hat{v}_i(t)) \\ &+ c_1((\hat{p}_0(t) - \hat{p}_i(t))) + c_2((\hat{v}_0(t) - \hat{v}_i(t))) \end{aligned} \quad (16)$$

where $\hat{x}_0(t)$ is the state estimation of a real leader or a virtual leader indexed by number zero, which can be regarded as consensus reference. If the agent i is informed by the leader, $a_{i0} = 1$, otherwise, $a_{i0} = 0$. $\hat{f}_i(t)$ is the estimation of formation reference at time $t \in [t_k^i, t_{k+1}^i)$ [83]. As for leaderless consensus Problem 1 and formation control without a reference leader in Problem 3, the terms related to a_{i0} are zero [51], [52]. For flocking control in (16), the collaborative error composes of the velocity consensus error and flocking tracking error, where $\hat{x}_0(t) = [\hat{p}_0(t), \hat{v}_0(t)]$ is the estimation of a flocking reference trajectory with position and velocity [103].

2) *Control Strategies*: are also designed in terms of intelligent control Problems 1–5. For consensus and formation control problems, the controller aims to eliminate collaborative errors by

$$u_i(t) = K_i \phi_i(t) \quad (17)$$

where K_i is the control law matrix based on ARE (algebraic Riccati equation) [9], [47], [49] and LMI (linear matrix inequality) [13], [48], [83]. Quantized event-triggered control is another improvement direction to save more network resources under limited bandwidth by the quantized control law

$$u_i(t) = K_i q_u(\phi_i(t)) \quad (18)$$

where q_u is a quantized function to convert the collaborative error into a discrete form [48], [53].

For flocking control in Problem 5, one form of control protocol is

$$u_i(t) = \sum_{j \in N_i} \psi(\|\hat{p}_j - \hat{p}_i\|_\sigma) (\mathbf{n}_{ij}) + \phi_i(t) \quad (19)$$

where ψ is a potential function [101]–[103] to remain close to neighbors without collisions. Another control strategy is based on the distributed model predictive control by solving optimization problems to satisfy the flocking rules [104], [106].

3) *Proof That the Zeno Phenomenon Is Excluded*: If an event is triggered infinitely within a finite time, the phenomenon is called Zeno phenomenon [122]. In the study of an event-based mechanism, one of the key tasks is to exclude Zeno phenomenon. One widely used method is to prove that there must be a positive lower bound on the interval length between any two trigger moments [57], [105], [106]

$$t_{k+1}^i - t_k^i \geq \tau > 0 \quad (20)$$

where τ is a positive constant that ensures the Zeno phenomenon is excluded.

The work in [123] also looked at the issue, and proposed another method. That is, if Zeno's behavior is assumed to exist, then there is at least a gathering point for the time-triggered sequence. By verifying that this assumption contradicts the existing attributes of the system, it is proved that Zeno's behavior can be excluded.

In addition to deal with the limited bandwidth of communication networks, event-based strategies are used to solve other problems subject to time delay [55], [56], [123], network attacks [22], [57], [74], switching topologies [9], and multiplicative faults [47]. However, to the best of our knowledge, there is a lack of fully distributed trigger strategies and methods to effectively exclude Zeno behavior, which does not involve any global information (such as the number of agents, Laplace matrix, etc.), especially for generalized linear and complex nonlinear MASs.

C. Pinning-Based Intelligent Control

The realization of information collaboration for MASs usually requires the assumption of the original connectivity of

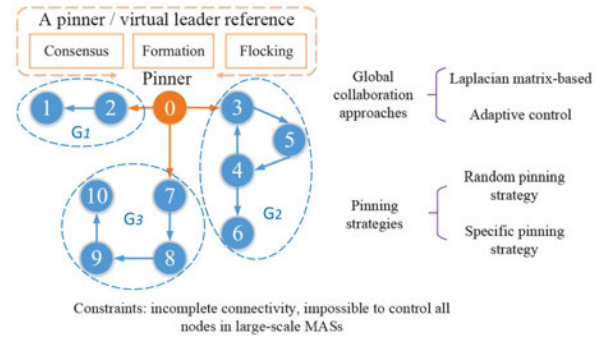


Fig. 9. Pinning-based intelligent control.

the communication topology [23]. However, this assumption is difficult to be satisfied especially for MASs under time-varying topologies. Moreover, for real-world large-scale MASs, such as complex power grids and multidrone light show systems, there are generally a large number of nodes or control points in their communication networks. It is usually difficult and expensive to put controllers to all nodes to control the whole system. Pinning-based control is one of the effective solutions for the issues. The basic idea of pinning control is by adding a pinner to control a small fraction of agents so that the whole MASs can achieve collaborative performance. As shown in Fig. 9, a pinner (or virtual leader) is added to the system and defines its desired trajectory, according to different intelligent control Problems 1–5. The pinner only controls some pinned agents (agent 2, agent 3, and agent 7) from different groups (G_1 , G_2 , and G_3) to ensure global collaboration, and the choice of pinned agent to inform the reference information is via the pinning strategies. Therefore, the realization of global collaboration and the design of pinning strategies are two key issues of the pinning-based intelligent control.

1) *Global Collaboration*: Pinning control is widely used to solve the synchronization of complex dynamic networks [58], [59], which is also a special case of leader–follower consensus problems defined in Problem 2 [53], [60]. The general first-order MASs composed of n agents under pinning control are modeled as

$$\begin{aligned} \dot{x}_i(t) = & f(x_i(t), t) + c \sum_{j=1}^n a_{ij} \Gamma(x_j(t) - x_i(t)) \\ & + cd_i \Gamma(s(t) - x_i(t)) \end{aligned} \quad (21)$$

where $x_i(t) \in \mathbb{R}^m$ and $f(x_i(t), t)$ are the state and nonlinear dynamic function of the i th agent, respectively. Coefficients c and d_i are the coupling strength and pinning control gain, respectively. $\Gamma \in \mathbb{R}^{n \times n}$ represents the inner coupling matrix. $s(t) \in \mathbb{R}^m$ denotes the state of a pinner modeled as

$$\dot{s}(t) = f(s(t), t)$$

where $f(s(t), t)$ is a nonlinear continuously differentiable function related to dynamic characteristics of the pinner [21]. One way to ensure consensus for whole MASs is to design suitable coupling strength, pinning control gain, and inner coupling matrix based on the Laplacian matrix, including the extension of the Laplacian matrix and the submatrix of

the Laplacian matrix (see [21] for details). However, this approach depends on the properties of the global communication matrix. The global information is difficult to obtain when the Laplacian matrix is time varying or stochastic due to the random selection of pinned agents. In order to overcome this drawback, adaptive pinning consensus strategies have been proposed for the first-order system without relying on any global information. At the same time, by introducing the adaptive rate in the coupling strength and control gain, the conservativeness of the Laplacian-based method has been reduced [21], [59].

In addition to the consensus problem, pinning-based control is also extended to solve formation tracking in Problem 4 [88] and flocking control in Problem 5 [107], [108]. A pinner or a virtual leader is used to provide the reference path for the agents to perform a collaborative task, such as the task of multidrone formation to reach a designated location with the desired trajectory. The relative formation reference respect for the virtual leader is also considered. The work in [88] provided a pinning-based control for nonlinear multidrone formation. Although a pinner describes the desired path of MASs, agents usually do not strictly follow the pinner when they encounter obstacles to avoid in the environment. In terms of the flocking problems defined in Problem 5, the pinner can be regarded as the flocking center with the desired trajectory. Pinning-based strategies for the flocking motion of a MAS have been developed under switching topologies [107] and sampled-data frameworks [108] to minimize the total cost considering pinner tracking, velocity consensus, and obstacle avoidance functions.

2) *Pinning Strategies*: They are investigated to determine the minimum number of nodes to be controlled and the specific pinned agents. Existing pinning strategies mainly include the random pinning strategy and the specific pinning strategy. The strategy of random selection first searches for strongly connected components [21] to assign groups of agents, and then randomly selects one agent in the group for control. Another method is to first arrange the nodes in descending order according to the difference between the out-degree and the in-degree, and select the first l nodes for control [60]. Both methods need to verify that the pinner is the root node, and there is at least one directed spanning tree in the entire MASs. It is difficult to ensure the connectivity of the original network, especially for arbitrary or changing topologies. Through a reasonable selection of pinned agents, the connectivity under the dynamic topology can be guaranteed at each moment. Many works have pointed out that the specific pinning strategy is more effective than the random pinning strategy in reducing the number of pinned agents [58], [60], [107].

Pinning-based control has been also introduced into some methodologies in control theory, such as the impulsive control [124], robust H_∞ control [125] and finite-time control [126] to improve the system performance under time delay and disturbances.

D. Resilient Control

MASs are likely to suffer from malicious attacks and corruption of sensory data or manipulation of actuators inputs,

which can severely and adversely affect system performance. For example, in a denial-of-service attack (DoS), the attacker intends to deny access to the data by making it unavailable to systems. Recently, considerable efforts have been made based on resilient control to detect and defend against attacks for MASs in terms of consensus, formation control, and flocking control defined in Problems 1–5.

A novel event-based resilient control was proposed in [61], which controlled the input signal rather than the state measurement error to solve leaderless consensus in Problem 1 under DoS attacks. For leader–follower consensus in Problem 2 under DoS attacks, the work in [62] provided a distributed fixed-time observer and an improved resilient observer to accurately estimate the leader's information, thus eliminating or weakening the influence of DoS. Reliable formation tracking control for MAS under quantized communication and false data injection (FDI) attacks was investigated in [89] based on a distributed filter with adaptive attack compensator. Both the system reliability in the attacked case and original performance in a no-attack case can be guaranteed with the developed filter. It can also achieve cooperative output regulation of MAS when the communication is not quantized but with potential attacks. For the unbounded malicious attacks, a fully distributed attack-resilient control protocol was proposed in [90] to solve the time-varying formation tracking problem defined in Problem 4. The bounded system stability and uniformly ultimately bounded synchronization performance have been guaranteed. Considering the presence of noncooperative robots, Saulnier *et al.* [109] developed a resilient flocking control approach for Problem 5. The proposed dynamic connectivity management and switching control strategies restricted the communication topology within the resilient threshold and allowed the mobile robots to achieve consensus along with the motion.

To the best of our knowledge, there are still a lack of effective detection and defence theoretical frameworks for MASs under multiple attacks.

E. Intelligent Control for Homogeneous MASs

MASs can be divided into homogeneous and heterogeneous systems, depending on whether the system dynamics are the same or not. In addition to the limitations of the interaction level, the constraints of the system level include nonlinear dynamics, heterogeneous dynamics, system uncertainties, external interference, and actuator and sensor failures.

For linear homogeneous MASs, there are a large number of works on consensus Problems 1 and 2 [38], [39], [63], formation control Problems 3 and 4 [91], [92], and flocking control Problem 5 [110], [111]. However, in practical applications, an agent is always subject to nonlinear dynamics, such as the flight control for multidrone formation [88] and flocking control for robots [102]. Various nonlinear control approaches have been developed for nonlinear MASs under uncertainties and bounded external disturbances, including adaptive control [64], backstepping scheme [88], sliding mode control [7], [93], neural network [27], [93], and fuzzy control [115]. Sliding mode control has been widely used to

control nonlinear systems with uncertainties and unknown disturbances, as the controllers can be designed to compensate for the uncertainties and disturbances.

Unforeseen threats may occur in system components, such as sensors, actuators, and controllers. Passive fault-tolerant control is one of the popular methods for managing physical faults or damages, with adaptive strategies. The work in [127] investigated a H_∞ consensus protocol together with an adaptive compensator to tolerate sensor and actuator faults for Problem 1. In [128], a cooperative adaptive fault-tolerant fuzzy control was proposed to solve leader–follower consensus Problem 2 of networked MASs with time-varying actuator faults. Adaptive formation control laws for Problem 4 were designed for unmanned aerial vehicles (UAVs) to tolerate actuator faults [129]. As a typical method of fault-tolerant control, adaptive controllers are used for compensating physical faults and passively tolerate system failures without changing the controller structure. However, the fixed structure has conservative problems and cannot optimize the system performance. To overcome the shortcomings, active fault-tolerant control strategies have been proposed for MASs. The active strategies usually include two functions: 1) fault detection and 2) control reconfiguration [128], [130], [131]. However, most of the existing research results assume that the system will not diverge during the time period of fault detection. This assumption is limited to multiple failures and severe physical damage.

F. Intelligent Control for Heterogeneous MASs

During the past decades, research on intelligent control has gradually shifted from homogeneous MASs to heterogeneous MASs. One reason is that it is difficult to equip the same agents with all the necessary sensing and computing equipment. Even for the same agents with the same equipment, it is not truly homogeneous MASs due to asynchronous clocks and uncertainties. Most important of all, heterogeneous MASs have more advantages in terms of formation flexibility and complex task decomposition because of the different functions of individuals.

The primary research on heterogeneous MASs is focused on hybrid-order MASs, such as consensus Problem 1 of first-order and second-order hybrid MASs [66] and formation Problem 3 of first-order and fourth-order hybrid MASs [77]. Graph theory-based matrix methods and Lyapunov theory are used to solve intelligent problems of hybrid-order MASs, but the global communication matrix is required especially for high-order hybrid MASs. Another approach for hybrid-order MASs aims to convert the mixed-order MASs into the same-order MASs and achieve the state consensus of the corresponding order by adding virtual zero states [30]. However, expanding the dimensions of agents may increase the computational load. State consensus is almost impossible for the general heterogeneous systems with different orders and different dynamics. Therefore, the output regulation for a single system is the focus, which can be extended to solve the output consensus problem of heterogeneous MASs [65]. Under the assumption that there exists a solution to the regulation equation [132],

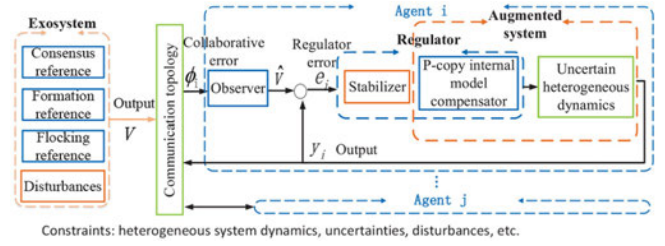


Fig. 10. Output regulation-based intelligent control for heterogeneous MASs.

the output consensus of heterogeneous MASs is solvable. Subsequently, the approach was promoted to solve the problems of formation control Problems 3 and 4 [79], [82], [133] and flocking control Problem 5 [118].

A general intelligent control framework of heterogeneous MASs is shown in Fig. 10, where the output of exosystem v includes reference information that needs to be tracked and disturbances need to be rejected. Since only a part of the agents can be informed by the exosystem, an observer is designed to estimate the output of the exosystem \hat{v} through collaborative error ϕ_i . The intelligent control problem of heterogeneous MASs is transformed into the stabilization problem of distributed augmented systems by a regulator. The regulator is composed of a stabilizer and p -copy internal model to stabilize the system and compensate for uncertainties, receptively. Augmented systems consist of a real agent and an internal model compensator. This method fundamentally uncouples the dynamics of heterogeneous agents and realizes the regulation of exosystem in a distributed manner.

A unified uncertain heterogeneous MAS with n agents of different orders can be modeled as

$$\begin{aligned} \dot{x}_i &= A_{wi}x_i + B_{wi}u_i + E_{wi}v \\ y_i &= C_{wi}x_i + D_{wi}u_i + F_{wi}v, \quad i = 1, 2, \dots, n \end{aligned} \quad (22)$$

where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}$, and $y_i \in \mathbb{R}^{p_i}$ represent the state, input, and output variables of the i th agent, respectively. Matrices $\star_{wi} = \star_i + \Delta\star_i$ are the corresponding system matrices with uncertainties $\Delta\star_i$, where $\star = A, B, C, D, E, F$.

The dynamic of an exosystem can be described as follows:

$$\dot{v} = Sv \quad (23)$$

where $v \in \mathbb{R}^{n_0}$ represent the state/output variables of the exosystem, which contains the reference information according to different intelligent Problems 1–5. For example, the dynamic of a time-varying formation [82] is considered as $\dot{f}_i = A_f^i f_i$, and the signal of exosystem v includes the information of virtual leader x_0 , time-varying formation dynamics f_i , and disturbance ω .

In order to achieve intelligent collaboration, the general stabilizer and internal model compensator are designed as follows:

$$\begin{aligned} u_i &= K_1^x x_i + K_1^z z_i \\ \dot{z}_i &= \Sigma_1 z_i + \Sigma_2 e_i \end{aligned} \quad (24)$$

where z_i is the state of the dynamic compensator. (Σ_1, Σ_2) is the p -copy internal model pair of S , based on the internal

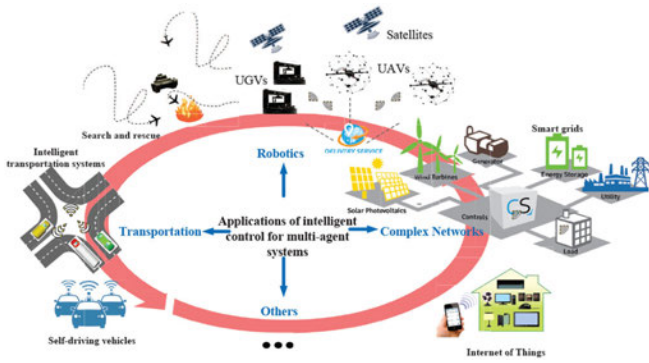


Fig. 11. Applications of intelligent control for MASs.

model principle [132]. The control law $K_i = (K_i^x, K_i^z)$ is obtained by the Lyapunov stability theory [82], [132]. The controller aims to eliminate regulation error, which is defined according to different intelligent control Problems 1–5. For example, for the leader–follower consensus in Problem 2, the error is defined as $e_i = y_i - \hat{v}$; for time-varying formation tracking in Problem 4 [79], [82], $e_i = y_i - \hat{v} - f_i$.

It should be noted that the observer converts the collaborative error ϕ_i to regulation error e_i based on the observation \hat{v} . Therefore, the design of observers is the core of decoupling heterogeneous dynamics and realizing distributed architecture. In [82], the adaptive distributed observer was proposed under the assumption that the exosystem dynamics S were globally known. Note that the design of the p -copy internal model system also required a known exosystem dynamic matrix. Model-based adaptive controls still involve global information, until an off-policy reinforcement learning strategy was investigated to make up the disadvantages [134].

The intelligent control for linear heterogeneous systems can also be extended to nonlinear systems to deal with consensus problems [30], [67], [68], formation control problems [79], [80], [94]–[96], and flocking control problems [120]. It is worth noting that this method is only applicable to leader-involved Problems 2, 4, and 5. For the leaderless problems as defined in Problems 1 and 3, it will raise a big challenge to discover the common internal model in heterogeneous systems in the absence of the exosystem. Learning-based approaches may be a possible way to address the problems.

IV. APPLICATIONS OF INTELLIGENT CONTROL FOR MASS

In this section, we present the achievements of intelligent control for MAS applications in the following directions: robotics, computer networks, transportation, and others. A summary of these applications is outlined in Fig. 11 and Table IV.

A. Robotics

The application of MASs on robotics has received extensive attention, especially on UAVs, unmanned ground vehicles (UGVs), and autonomous underwater vehicles (AUVs). The work in [135] proposed a modular architecture of multiUAV

TABLE IV
APPLICATIONS OF INTELLIGENT CONTROL FOR MASS

Applications	Feature-specific MAS	Refinance
Robotics	UAVs	[135]–[139]
	UGVs	[140]–[144]
	AUVs	[145], [146]
Complex networks	Smart grid systems	[147]–[149]
	Internet of Things	[150], [151]
Transportation	Traffic light systems	[152]–[154]
	Smart driving systems	[155]

collaboration system for search and rescue missions. The framework has been verified and evaluated by outdoor experiments of four prototype UAVs. A multiUAV collaborative approach in disaster management and civil security applications has been validated with real UAVs and wireless sensor networks [136]. Formation protocols and consensus approaches were used to achieve time-varying formations, which were tested and verified by distributed outdoor experiments with five quadrotors [138]. Relative information of neighboring UAVs can be used to construct a time-varying formation control protocol for swarm systems. An outdoor target enclosing experiment was carried out for three follower quadrotor UAVs to enclose a leader quadrotor UAV by time-varying formations [139]. In [140], a life support robot system was developed to perform domestic services that are useful to the well being of the elderly with walking disabilities. A frequent task during life support is the fetching of daily containers, such as serving drinks and food [140].

Furthermore, a large number of results on the formation control for hardware multivehicle platforms have been verified under laboratory conditions. A distributed formation control approach for multirotor UAVs was proposed and embedded into the onboard computational units to make them able to keep a balanced formation in 2-D and three dimensional environments [137]. The proposed formation control approach was proved to be feasible for arbitrary formation by both simulations and real-system experiment. In order to lead the UGVs moving into the desired formation quickly, a cooperative coevolutionary algorithm-based distributed model predictive control was proposed in [144], which can greatly improve the performance of formation control as performed by three mobile robots. An adaptive self-organizing map neural network was applied to keep the formation when agents move along the desired path [145]. Both simulations and real AUV systems demonstrated the fault-tolerant characteristic in obstacle avoidance and the benefit of balancing the workload and energy. A modified constrained adaptive controller was proposed to resolve the communication delay and actuator saturation [146]. Simulation and experimental validation showed that the method can effectively compensate for the effects of state delay in 2 and 5 s, respectively. Follower AUVs were able to follow the desired path within the accuracy of 5 cm.

B. Complex Networks

The resource-aware consensus theory of the MASs provides a theoretical reference for smart grid applications. In [147], the MAS-based algorithm has been applied to control voltage

and capacitor to optimally set the system. An event-triggered strategy is proved to be an effective method to reduce the communication burden of a network. The two-level reinforcement learning-based controller was proposed in [148], where the parameters were optimized by particle swarm. The results verified the feasibility of the proposed method. Within the decentralized system integrity protection set up, data-driven anomaly detection, and adaptive load rejection were studied in [149]. Anomaly detection has been converted to a multiclassification problem and can be performed by individual agents, but all the interconnected agents devoted to the final decision. Meanwhile, the proposed adaptive load rejection strategy can reduce the DoS attacks.

The application of MASs can be extended to the Internet of Things (IoT) where objects range from sensors to wearable devices. Agent-based resilient control plays a vital role in IoT networks. It is essential to effectively identify the malicious node and prevent further damage. A combined multiagent and multilayered game formulation was proposed in [150], which incorporated a trust model to assess the node/object. The proposed model can significantly improve the accuracy of intrusion detection by experimental test. IoT inevitably introduces a vast amount of real-time data. A multiagent-based real-time scheduling architecture was presented to optimally assign tasks according to the real-time status of machines [151].

C. Transportation

The large-scale intelligent transportation system is one of the typical applications of MASs. Taking the dimension, complicated dynamics, and uncertainties into consideration, Lin *et al.* [152] proposed a centralized multiagent control method with a serial framework. Agents communicate with their neighbors through a model-based predictive control method. Using the traffic data provided by the city of Toronto, an adaptive reinforcement learning-based traffic signal controller was proposed in [153], which can work in two modes: 1) decentralized and 2) centralized. However, the dynamic and complex traffic conditions make it difficult for the model-based and reinforcement learning-based models to make good decisions. In [154], a multiagent recurrent deep deterministic policy gradient algorithm was proposed to control the traffic light in land traffic. Decisions were made independently by each agent, thus avoiding the poor performance caused by an unstable environment. Autonomous driving is another application in intelligent control, among the key complex problems, the formation will be outstanding. Due to the formation changes with the traffic flow and conditions, a dynamic coordination graph was proposed to model the constantly changing topology to coordinate the maneuvers of grouped vehicles in [155], which was proved to be effective than some expert rules.

D. Others

The applications of MASs are not limited to the above-mentioned fields. They have also been widely applied to aerospace, agriculture, industrial production, and medical treatment, to name but a few.

V. CONCLUSION AND FUTURE RESEARCH CHALLENGES

In this article, we presented a survey of distributed intelligent control for MASs. Focusing on the constraints from the interaction level and system level, the recent results have been reviewed in terms of consensus problem, formation control problem, and flocking control problem. However, this is far from an exhaustive literature review and some important results might be missed due to the limitation of our knowledge. Furthermore, there still exist several challenges in this area deserving further study.

- 1) Security is highly challenging for MASs. Most existing works design resilient and robust strategies separately on interaction level and system level. For instance, distributed resilient control under attacks and communication problems tends to use network-level design, where the individual agent with high-fidelity dynamics is usually simplified, and fault-tolerance control mainly focuses on homogeneous system-level robustness. However, the separated security control design on two levels fails to realize quick stability and recover to optimal performance, which poses a threat to the survival of MASs under multiple threats and unknown environments. Therefore, high-reliability intelligent control under both two-level threats is still an open problem to be solved in the future.
- 2) The design of fully distributed intelligent control and its optimization is still considered as open issues. Although many studies have focused on distributed control approaches, some global information, such as the total number of the agents and the Laplace matrix of the communication topology are still being involved for high-order MASs in intelligent control designs. Verification on global stability, connectivity preservation under dynamic topology, proof of nonZeno phenomenon, and optimization of task assignments are usually not designed with a fully distributed framework. Adaptive control strategies and learning-based techniques are used to resolve this imperfect. However, they inevitably increase the computational load. In real-world applications, agents subject to limited computing capability need to perceive, make decisions, and take actions independently, which raises higher requirements for fully distributed algorithms and optimization techniques.
- 3) The research on intelligent control for heterogeneous MASs should be enhanced, especially in both theoretic research and applications of heterogeneous multivehicle systems. Although the works on heterogeneous MASs have received extensive attention in the past decades, most of them focused on the fundamental consensus problems and theoretic research. In fact, it is difficult to build truly homogeneous MASs in practical applications. The intelligent control of heterogeneous vehicles, such as UAVs, UGVs, and AUVs, is more promising in practical applications to achieve multidimensional collaboration under complementary capabilities.

- 4) Verifications for distributed intelligent control strategies in large-scale practical application scenarios are urgently needed. Most existing results are obtained under laboratory conditions with centralized structures. For large-scale MASs, very few studies are carried out in the actual application environment, which leads to the urgent requirements of verifications in the actual application environment.

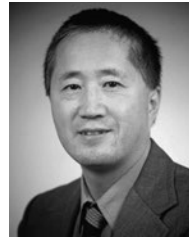
REFERENCES

- [1] M. Wooldridge, *An Introduction to Multiagent Systems*. Hoboken, NJ, USA: Wiley, 2009.
- [2] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 427–438, Feb. 2013.
- [3] J. Qin, Q. Ma, Y. Shi, and L. Wang, "Recent advances in consensus of multi-agent systems: A brief survey," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4972–4983, Jun. 2017.
- [4] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: A review from the swarm engineering perspective," *Swarm Intell.*, vol. 7, no. 1, pp. 1–41, 2013.
- [5] J. Kennedy, "Swarm intelligence," in *Handbook of Nature-Inspired and Innovative Computing*. Boston, MA, USA: Springer, 2006, pp. 187–219.
- [6] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, Mar. 2015.
- [7] Q. Shen, P. Shi, and Y. Shi, "Distributed adaptive fuzzy control for nonlinear multiagent systems via sliding mode observers," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3086–3097, Dec. 2016.
- [8] P. Shi and Q. Shen, "Observer-based leader-following consensus of uncertain nonlinear multi-agent systems," *Int. J. Robust Nonlinear Control*, vol. 27, no. 17, pp. 3794–3811, 2017.
- [9] Z.-G. Wu, Y. Xu, R. Lu, Y. Wu, and T. Huang, "Event-triggered control for consensus of multiagent systems with fixed/switching topologies," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 10, pp. 1736–1746, Oct. 2018.
- [10] S. Rao and D. Ghose, "Sliding mode control-based autopilots for leaderless consensus of unmanned aerial vehicles," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 5, pp. 1964–1972, Sep. 2014.
- [11] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.
- [12] Z. Han, K. Guo, L. Xie, and Z. Lin, "Integrated relative localization and leader–follower formation control," *IEEE Trans. Autom. Control*, vol. 64, no. 1, pp. 20–34, Jan. 2019.
- [13] X. Li, X. Dong, Q. Li, and Z. Ren, "Event-triggered time-varying formation control for general linear multi-agent systems," *J. Franklin Inst.*, vol. 356, no. 17, pp. 10179–10195, 2019.
- [14] K. D. Do, "Flocking for multiple elliptical agents with limited communication ranges," *IEEE Trans. Robot.*, vol. 27, no. 5, pp. 931–942, Oct. 2011.
- [15] A. Weinstein, A. Cho, G. Loianno, and V. Kumar, "Visual inertial odometry swarm: An autonomous swarm of vision-based quadrotors," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1801–1807, Jul. 2018.
- [16] F. Schilling, J. Lecoeur, F. Schiano, and D. Floreano, "Learning vision-based flight in drone swarms by imitation," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 4523–4530, Oct. 2019.
- [17] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader–follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, 2008.
- [18] L. Ding, Q.-L. Han, and G. Guo, "Network-based leader-following consensus for distributed multi-agent systems," *Automatica*, vol. 49, no. 7, pp. 2281–2286, 2013.
- [19] H. Yu, P. Shi, and C. C. Lim, "Robot formation control in stealth mode with scalable team size," *Int. J. Control*, vol. 89, no. 11, pp. 2155–2168, 2016.
- [20] D. V. Dimarogonas, E. Frazzoli, and K. H. Johansson, "Distributed event-triggered control for multi-agent systems," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1291–1297, May 2012.
- [21] W. Xing, P. Shi, R. K. Agarwal, and Y. Zhao, "A survey on global pinning synchronization of complex networks," *J. Franklin Inst.*, vol. 356, no. 6, pp. 3590–3611, 2019.
- [22] D. Zhang, Y. Tang, Z. Ding, and F. Qian, "Event-based resilient formation control of multiagent systems," *IEEE Trans. Cybern.*, early access, Apr. 26, 2019, doi: [10.1109/TCYB.2019.2910614](https://doi.org/10.1109/TCYB.2019.2910614).
- [23] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.
- [24] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
- [25] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.
- [26] G. Song, P. Shi, S. Wang, and J.-S. Pan, "A new finite-time cooperative control algorithm for uncertain multi-agent systems," *Int. J. Syst. Sci.*, vol. 50, no. 5, pp. 1006–1016, 2019.
- [27] J. Ni and P. Shi, "Adaptive neural network fixed-time leader–follower consensus for multiagent systems with constraints and disturbances," *IEEE Trans. Cybern.*, early access, Feb. 24, 2020, doi: [10.1109/TCYB.2020.2967995](https://doi.org/10.1109/TCYB.2020.2967995).
- [28] P. Shi and J. Yu, "Dissipativity-based consensus for fuzzy multi-agent systems under switching directed topologies," *IEEE Trans. Fuzzy Syst.*, early access, Jan. 27, 2020, doi: [10.1109/TFUZZ.2020.2969391](https://doi.org/10.1109/TFUZZ.2020.2969391).
- [29] L. Ji, T. Gao, and X. Liao, "Couple-group consensus for cooperative-competitive heterogeneous multiagent systems: Hybrid adaptive and pinning methods," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Nov. 15, 2019, doi: [10.1109/TSMC.2019.2951787](https://doi.org/10.1109/TSMC.2019.2951787).
- [30] X. Li, P. Shi, and Y. Wang, "Distributed cooperative adaptive tracking control for heterogeneous systems with hybrid nonlinear dynamics," *Nonlinear Dyn.*, vol. 95, no. 3, pp. 2131–2141, 2019.
- [31] X. Li and P. Shi, "Cooperative fault-tolerant tracking control of heterogeneous hybrid-order mechanical systems with actuator and amplifier faults," *Nonlinear Dyn.*, vol. 98, no. 1, pp. 447–462, 2019.
- [32] B. Yan, C. Wu, and P. Shi, "Formation consensus for discrete-time heterogeneous multi-agent systems with link failures and actuator/sensor faults," *J. Franklin Inst.*, vol. 356, pp. 6547–6570, Aug. 2019.
- [33] Y. Su and J. Huang, "Cooperative robust output regulation of a class of heterogeneous linear uncertain multi-agent systems," *Int. J. Robust Nonlinear Control*, vol. 24, no. 17, pp. 2819–2839, 2014.
- [34] Y. Yang, D. Yue, and C. Dou, "Distributed adaptive output consensus control of a class of heterogeneous multi-agent systems under switching directed topologies," *Inf. Sci.*, vol. 345, pp. 294–312, Jun. 2016.
- [35] Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and Y. Shen, "Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control," *Automatica*, vol. 93, pp. 26–32, Jul. 2018.
- [36] S. Zheng, P. Shi, S. Wang, and Y. Shi, "Adaptive neural control for a class of nonlinear multiagent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 30, 2020, doi: [10.1109/TNNLS.2020.2979266](https://doi.org/10.1109/TNNLS.2020.2979266).
- [37] B. Bollobás, *Modern Graph Theory*, vol. 184. Berlin, Germany: Springer, 2013.
- [38] B. Cheng and Z. Li, "Fully distributed event-triggered protocols for linear multiagent networks," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1655–1662, Apr. 2019.
- [39] R. Sakthivel, R. Sakthivel, B. Kaviarasan, H. Lee, and Y. Lim, "Finite-time leaderless consensus of uncertain multi-agent systems against time-varying actuator faults," *Neurocomputing*, vol. 325, pp. 159–171, Jan. 2019.
- [40] J. Huang, W. Wang, C. Wen, J. Zhou, and G. Li, "Distributed adaptive leader–follower and leaderless consensus control of a class of strict-feedback nonlinear systems: A unified approach," *Automatica*, vol. 118, Aug. 2020, Art. no. 109021.
- [41] Y. Lv, G. Wen, T. Huang, and Z. Duan, "Adaptive attack-free protocol for consensus tracking with pure relative output information," *Automatica*, vol. 117, Jul. 2020, Art. no. 108998.
- [42] M. Lu and J. Huang, "Internal model approach to cooperative robust output regulation for linear uncertain time-delay multiagent systems," *Int. J. Robust Nonlinear Control*, vol. 28, no. 6, pp. 2528–2542, 2018.
- [43] Q. Shen, P. Shi, J. Zhu, S. Wang, and Y. Shi, "Neural networks-based distributed adaptive control of nonlinear multiagent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 3, pp. 1010–1021, Mar. 2020.
- [44] Q. Song, F. Liu, J. Cao, and W. Yu, "M-matrix strategies for pinning-controlled leader-following consensus in multiagent systems with nonlinear dynamics," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1688–1697, Dec. 2013.

- [45] Q. Ma, Z. Wang, and G. Miao, "Second-order group consensus for multi-agent systems via pinning leader-following approach," *J. Franklin Inst.*, vol. 351, no. 3, pp. 1288–1300, 2014.
- [46] L. Li, P. Shi, and C. K. Ahn, "Distributed iterative FIR consensus filter for multiagent systems over sensor networks," *IEEE Trans. Cybern.*, early access, Dec. 9, 2020, doi: [10.1109/TCYB.2020.3035866](https://doi.org/10.1109/TCYB.2020.3035866).
- [47] D. Ye, M.-M. Chen, and H.-J. Yang, "Distributed adaptive event-triggered fault-tolerant consensus of multiagent systems with general linear dynamics," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 757–767, Mar. 2019.
- [48] C. Gong, G. Zhu, and P. Shi, "Adaptive event-triggered and double-quantized consensus of leader-follower multiagent systems with semi-Markovian jump parameters," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Dec. 20, 2019, doi: [10.1109/TSMC.2019.2957530](https://doi.org/10.1109/TSMC.2019.2957530).
- [49] R. Yang, H. Zhang, G. Feng, H. Yan, and Z. Wang, "Robust cooperative output regulation of multi-agent systems via adaptive event-triggered control," *Automatica*, vol. 102, pp. 129–136, Apr. 2019.
- [50] H. Zhang, J. Chen, Z. Wang, C. Fu, and S. Song, "Distributed event-triggered control for cooperative output regulation of multiagent systems with an online estimation algorithm," *IEEE Trans. Cybern.*, early access, Jun. 8, 2020, doi: [10.1109/TCYB.2020.2991761](https://doi.org/10.1109/TCYB.2020.2991761).
- [51] B. Hu, Z.-H. Guan, and M. Fu, "Distributed event-driven control for finite-time consensus," *Automatica*, vol. 103, pp. 88–95, May 2019.
- [52] W. Zhu, Z.-P. Jiang, and G. Feng, "Event-based consensus of multi-agent systems with general linear models," *Automatica*, vol. 50, no. 2, pp. 552–558, 2014.
- [53] Z. Wu, Y. Xu, Y. Pan, P. Shi, and Q. Wang, "Event-triggered pinning control for consensus of multiagent systems with quantized information," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 11, pp. 1929–1938, Nov. 2018.
- [54] W. Zou, P. Shi, Z. Xiang, and Y. Shi, "Consensus tracking control of switched stochastic nonlinear multiagent systems via event-triggered strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 3, pp. 1036–1045, Mar. 2020.
- [55] C. Deng, M. J. Er, G.-H. Yang, and N. Wang, "Event-triggered consensus of linear multiagent systems with time-varying communication delays," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 2916–2925, Jul. 2020.
- [56] X. Zhou, P. Shi, C.-C. Lim, C. Yang, and W. Gui, "Event based guaranteed cost consensus for distributed multi-agent systems," *J. Franklin Inst.*, vol. 352, no. 9, pp. 3546–3563, 2015.
- [57] Y. Xu, M. Fang, P. Shi, and Z. Wu, "Event-based secure consensus of multiagent systems against DoS attacks," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3468–3476, Aug. 2020.
- [58] X. F. Wang and G. Chen, "Pinning control of scale-free dynamical networks," *Physica A, Stat. Mech. Appl.*, vol. 310, nos. 3–4, pp. 521–531, 2002.
- [59] H. Chen, P. Shi, and C. Lim, "Cluster synchronization for neutral stochastic delay networks via intermittent adaptive control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3246–3259, Nov. 2019.
- [60] Q. Song, J. Cao, and W. Yu, "Second-order leader-following consensus of nonlinear multi-agent systems via pinning control," *Syst. Control Lett.*, vol. 59, no. 9, pp. 553–562, Sep. 2010.
- [61] Y. Xu, M. Fang, Z. Wu, Y. Pan, M. Chadli, and T. Huang, "Input-based event-triggering consensus of multiagent systems under denial-of-service attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 4, pp. 1455–1464, Apr. 2020.
- [62] H. Yang and D. Ye, "Observer-based fixed-time secure tracking consensus for networked high-order multiagent systems against DoS attacks," *IEEE Trans. Cybern.*, early access, Jul. 22, 2020, doi: [10.1109/TCYB.2020.3005354](https://doi.org/10.1109/TCYB.2020.3005354).
- [63] X. Ai, S. Song, and K. You, "Second-order consensus of multi-agent systems under limited interaction ranges," *Automatica*, vol. 68, pp. 329–333, Jun. 2016.
- [64] Q. Shen, P. Shi, J. Zhu, and L. Zhang, "Adaptive consensus control of leader-following systems with transmission nonlinearities," *Int. J. Control*, vol. 92, no. 2, pp. 317–328, 2019.
- [65] Y. Su and J. Huang, "Cooperative output regulation of linear multi-agent systems," *IEEE Trans. Autom. Control*, vol. 57, no. 4, pp. 1062–1066, Apr. 2012.
- [66] S.-L. Du, W. Xia, X.-M. Sun, and W. Wang, "Sampled-data-based consensus and L_2 -gain analysis for heterogeneous multiagent systems," *IEEE Trans. Cybern.*, vol. 47, no. 6, pp. 1523–1531, Jun. 2017.
- [67] W. Zhang, D. W. Ho, Y. Tang, and Y. Liu, "Quasi-consensus of heterogeneous-switched nonlinear multiagent systems," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3136–3146, Jul. 2020.
- [68] S. Li, M. J. Er, and J. Zhang, "Distributed adaptive fuzzy control for output consensus of heterogeneous stochastic nonlinear multiagent systems," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1138–1152, Jun. 2018.
- [69] W. Zou, C. K. Ahn, and Z. Xiang, "Fuzzy-approximation-based distributed fault-tolerant consensus for heterogeneous switched nonlinear multiagent systems," *IEEE Trans. Fuzzy Syst.*, early access, Jul. 16, 2020, doi: [10.1109/TFUZZ.2020.3009730](https://doi.org/10.1109/TFUZZ.2020.3009730).
- [70] W. Ren, "Consensus strategies for cooperative control of vehicle formations," *IET Control Theory Appl.*, vol. 1, no. 2, pp. 505–512, Mar. 2007.
- [71] S. He, M. Wang, S.-L. Dai, and F. Luo, "Leader-follower formation control of USVs with prescribed performance and collision avoidance," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 572–581, Jan. 2019.
- [72] G. Lee and D. Chwa, "Decentralized behavior-based formation control of multiple robots considering obstacle avoidance," *Intell. Serv. Robot.*, vol. 11, no. 1, pp. 127–138, Jan. 2018.
- [73] D. Zhou, Z. Wang, and M. Schwager, "Agile coordination and assistive collision avoidance for quadrotor swarms using virtual structures," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 916–923, Aug. 2018.
- [74] Y. Tang, D. Zhang, P. Shi, W. Zhang, and F. Qian, "Event-based formation control for multi-agent systems under DoS attacks," *IEEE Trans. Autom. Control*, early access, Mar. 10, 2020, doi: [10.1109/TAC.2020.2979936](https://doi.org/10.1109/TAC.2020.2979936).
- [75] H. Yu, P. Shi, and C. C. Lim, "Scalable formation control in stealth with limited sensing range," *Int. J. Robust Nonlinear Control*, vol. 27, no. 3, pp. 410–433, 2017.
- [76] Y. Liu, H. Yu, P. Shi, and C.-C. Lim, "Formation control and collision avoidance for a class of multi-agent systems," *J. Franklin Inst.*, vol. 356, no. 10, pp. 5395–5420, 2019.
- [77] B. Yan, P. Shi, C.-C. Lim, C. Wu, and Z. Shi, "Optimally distributed formation control with obstacle avoidance for mixed-order multi-agent systems under switching topologies," *IET Control Theory Appl.*, vol. 12, no. 13, pp. 1853–1863, Sep. 2018.
- [78] X. Dong, Y. Hua, Y. Zhou, Z. Ren, and Y. Zhong, "Theory and experiment on formation-containment control of multiple multirotor unmanned aerial vehicle systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 1, pp. 229–240, Jan. 2019.
- [79] S. Li, J. Zhang, X. Li, F. Wang, X. Luo, and X. Guan, "Formation control of heterogeneous discrete-time nonlinear multi-agent systems with uncertainties," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4730–4740, Jun. 2017.
- [80] C. Yuan, S. Licht, and H. He, "Formation learning control of multiple autonomous underwater vehicles with heterogeneous nonlinear uncertain dynamics," *IEEE Trans. Cybern.*, vol. 48, no. 10, pp. 2920–2934, Oct. 2018.
- [81] C. Deng and W.-W. Che, "Fault-tolerant fuzzy formation control for a class of nonlinear multiagent systems under directed and switching topology," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Dec. 5, 2019, doi: [10.1109/TSMC.2019.2954870](https://doi.org/10.1109/TSMC.2019.2954870).
- [82] W. Jiang, G. Wen, Z. Peng, T. Huang, and A. Rahmani, "Fully distributed formation-containment control of heterogeneous linear multiagent systems," *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3889–3896, Sep. 2019.
- [83] X. Ge and Q.-L. Han, "Distributed formation control of networked multi-agent systems using a dynamic event-triggered communication mechanism," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8118–8127, Oct. 2017.
- [84] Y. Liu, P. Shi, and C. C. Lim, "Collision-free formation control for multi-agent systems with dynamic mapping," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 10, pp. 1984–1988, Oct. 2020.
- [85] H. Yu, P. Shi, C.-C. Lim, and Y. Liu, "Probability-triggered formation control with adaptive roles," *Int. J. Control*, vol. 93, no. 8, pp. 1989–2000, 2020.
- [86] Y. Liu, P. Shi, H. Yu, and C. C. Lim, "Event-triggered probability-driven adaptive formation control for multiple elliptical agents," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Oct. 16, 2020, doi: [10.1109/TSMC.2020.3026029](https://doi.org/10.1109/TSMC.2020.3026029).
- [87] Q. Xiao, F. L. Lewis, and Z. Zeng, "Event-based time-interval pinning control for complex networks on time scales and applications," *IEEE Trans. Ind. Electron.*, vol. 65, no. 11, pp. 8797–8808, Nov. 2018.
- [88] Y. Kartal, K. Subbarao, N. R. Gans, A. Dogan, and F. Lewis, "Distributed backstepping based control of multiple UAV formation flight subject to time delays," *IET Control Theory Appl.*, vol. 14, no. 12, pp. 1628–1638, Aug. 2020.

- [89] X. Huang and J. Dong, "Reliable leader-to-follower formation control of multiagent systems under communication quantization and attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 1, pp. 89–99, Jan. 2020.
- [90] S. Zuo and D. Yue, "Resilient output formation containment of heterogeneous multigroup systems against unbounded attacks," *IEEE Trans. Cybern.*, early access, Jun. 30, 2020, doi: [10.1109/TCYB.2020.2998333](https://doi.org/10.1109/TCYB.2020.2998333).
- [91] X. Liu, Z. Ji, T. Hou, and H. Yu, "Decentralized stabilizability and formation control of multi-agent systems with antagonistic interactions," *ISA Trans.*, vol. 89, pp. 58–66, Jun. 2019.
- [92] M. Yu, H. Wang, G. Xie, and K. Jin, "Event-triggered circle formation control for second-order-agent system," *Neurocomputing*, vol. 275, pp. 462–469, Jan. 2018.
- [93] Y. Fei, P. Shi, and C.-C. Lim, "Neural network adaptive dynamic sliding mode formation control of multi-agent systems," *Int. J. Syst. Sci.*, vol. 51, no. 11, pp. 2025–2040, 2020.
- [94] R. Rahimi, F. Abdollahi, and K. Naqshi, "Time-varying formation control of a collaborative heterogeneous multi agent system," *Robot. Autom. Syst.*, vol. 62, no. 12, pp. 1799–1805, Dec. 2014.
- [95] J. Yu, X. Dong, Q. Li, and Z. Ren, "Practical time-varying formation tracking for second-order nonlinear multiagent systems with multiple leaders using adaptive neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6015–6025, Dec. 2018.
- [96] A. R. Mehrabian and K. Khorasani, "Distributed formation recovery control of heterogeneous multiagent Euler–Lagrange systems subject to network switching and diagnostic imperfections," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 6, pp. 2158–2166, Nov. 2016.
- [97] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *Proc. ACM SIGGRAPH Conf. Comput. Graph.*, Jul. 1987, pp. 25–34.
- [98] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, no. 6, p. 1226, Aug. 1995.
- [99] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 863–868, May 2007.
- [100] H. Yu, C.-C. Lim, R. Hunjet, and P. Shi, "Flocking and topology manipulation based on space partitioning," *Robot. Autom. Syst.*, vol. 124, Feb. 2020, Art. no. 103328, doi: [10.1016/j.robot.2019.103328](https://doi.org/10.1016/j.robot.2019.103328).
- [101] F. Sun, R. Wang, W. Zhu, and Y. Li, "Flocking in nonlinear multi-agent systems with time-varying delay via event-triggered control," *Appl. Math. Comput.*, vol. 350, pp. 66–77, Jun. 2019.
- [102] Y. Shen, Z. Kong, and L. Ding, "Flocking of multi-agent system with nonlinear dynamics via distributed event-triggered control," *Appl. Sci.*, vol. 9, no. 7, p. 1336, Jan. 2019.
- [103] P. Yu, L. Ding, Z.-W. Liu, and Z.-H. Guan, "Leader–follower flocking based on distributed event-triggered hybrid control," *Int. J. Robust Nonlinear Control*, vol. 26, no. 1, pp. 143–153, 2016.
- [104] Y. Hu, J. Zhan, and X. Li, "Self-triggered distributed model predictive control for flocking of multi-agent systems," *IET Control Theory Appl.*, vol. 12, no. 18, pp. 2441–2448, Dec. 2018.
- [105] Y. Xu, M. Fang, P. Shi, Y. Pan, and C. K. Ahn, "Multileader multiagent systems containment control with event-triggering," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Mar. 12, 2019, doi: [10.1109/TSMC.2019.2899967](https://doi.org/10.1109/TSMC.2019.2899967).
- [106] L. Li, P. Shi, R. K. Agarwal, C. K. Ahn, and W. Xing, "Event-triggered model predictive control for multiagent systems with communication constraints," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Aug. 21, 2019, doi: [10.1109/TSMC.2019.2932838](https://doi.org/10.1109/TSMC.2019.2932838).
- [107] J. Gao, X. Xu, N. Ding, and E. Li, "Flocking motion of multi-agent system by dynamic pinning control," *IET Control Theory Appl.*, vol. 11, no. 5, pp. 714–722, Mar. 2017.
- [108] S. Yazdani, M. Haeri, and H. Su, "Sampled-data leader–follower algorithm for flocking of multi-agent systems," *IET Control Theory Appl.*, vol. 13, no. 5, pp. 609–619, Mar. 2019.
- [109] K. Saulnier, D. Saldaña, A. Prorok, G. J. Pappas, and V. Kumar, "Resilient flocking for mobile robot teams," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1039–1046, Apr. 2017.
- [110] S. H. Semnani and O. A. Basir, "Semi-flocking algorithm for motion control of mobile sensors in large-scale surveillance systems," *IEEE Trans. Cybern.*, vol. 45, no. 1, pp. 129–137, Jan. 2015.
- [111] D. Lee and M. W. Spong, "Stable flocking of multiple inertial agents on balanced graphs," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1469–1475, Aug. 2007.
- [112] W. Zhao, H. Chu, M. Zhang, T. Sun, and L. Guo, "Flocking control of fixed-wing UAVs with cooperative obstacle avoidance capability," *IEEE Access*, vol. 7, pp. 17798–17808, 2019.
- [113] T. Han and S. S. Ge, "Styled-velocity flocking of autonomous vehicles: A systematic design," *IEEE Trans. Autom. Control*, vol. 60, no. 8, pp. 2015–2030, Aug. 2015.
- [114] T. Ibuki, S. Wilson, J. Yamauchi, M. Fujita, and M. Egerstedt, "Optimization-based distributed flocking control for multiple rigid bodies," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1891–1898, Apr. 2020.
- [115] B. K. Sahu and B. Subudhi, "Flocking control of multiple AUVs based on fuzzy potential functions," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 2539–2551, Oct. 2018.
- [116] D. Gu and Z. Wang, "Leader–follower flocking: Algorithms and experiments," *IEEE Trans. Control Syst. Technol.*, vol. 17, no. 5, pp. 1211–1219, Sep. 2009.
- [117] Y. Jia and L. Wang, "Leader–follower flocking of multiple robotic fish," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 3, pp. 1372–1383, Jun. 2015.
- [118] A. Prorok, M. A. Hsieh, and V. Kumar, "The impact of diversity on optimal control policies for heterogeneous robot swarms," *IEEE Trans. Robot.*, vol. 33, no. 2, pp. 346–358, Apr. 2017.
- [119] S. Chen, H. Pei, Q. Lai, and H. Yan, "Multitarget tracking control for coupled heterogeneous inertial agents systems based on flocking behavior," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 12, pp. 2605–2611, Dec. 2019.
- [120] Q. Zhang, Y. Hao, Z. Yang, and Z. Chen, "Adaptive flocking of heterogeneous multi-agents systems with nonlinear dynamics," *Neurocomputing*, vol. 216, pp. 72–77, Dec. 2016.
- [121] H. Yu, P. Shi, C.-C. Lim, and D. Wang, "Formation control for multi-robot systems with collision avoidance," *Int. J. Control*, vol. 92, no. 10, pp. 2223–2234, 2018.
- [122] A. D. Ames, A. Abate, and S. Sastry, "Sufficient conditions for the existence of zeno behavior," in *Proc. IEEE 44th Conf. Decis. Control*, Seville, Spain, Dec. 2005, pp. 696–701.
- [123] W. Zhu and Z. Jiang, "Event-based leader-following consensus of multi-agent systems with input time delay," *IEEE Trans. Autom. Control*, vol. 60, no. 5, pp. 1362–1367, May 2015.
- [124] H. Chen, P. Shi, and C. Lim, "Synchronization control for neutral stochastic delay Markov networks via single pinning impulsive strategy," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 12, pp. 5406–5419, Dec. 2020, doi: [10.1109/TSMC.2018.2882836](https://doi.org/10.1109/TSMC.2018.2882836).
- [125] W. Xing, P. Shi, R. K. Agarwal, and L. Li, "Robust H_∞ pinning synchronization for complex networks with event-triggered communication scheme," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 12, pp. 5233–5245, Dec. 2020, doi: [10.1109/TCSL.2020.3004170](https://doi.org/10.1109/TCSL.2020.3004170).
- [126] Z. Tang, J. H. Park, and H. Shen, "Finite-time cluster synchronization of Lur'e networks: A nonsmooth approach," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 8, pp. 1213–1224, Aug. 2018.
- [127] H. Modares, B. Kiumarsi, F. L. Lewis, F. Ferrese, and A. Davoudi, "Resilient and robust synchronization of multiagent systems under attacks on sensors and actuators," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1240–1250, Mar. 2020.
- [128] Q. Shen, B. Jiang, P. Shi, and J. Zhao, "Cooperative adaptive fuzzy tracking control for networked unknown nonlinear multiagent systems with time-varying actuator faults," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 3, pp. 494–504, Jun. 2014.
- [129] X. Yu, Z. Liu, and Y. Zhang, "Fault-tolerant formation control of multiple UAVs in the presence of actuator faults," *Int. J. Robust Nonlinear Control*, vol. 26, no. 12, pp. 2668–2685, 2016.
- [130] F. Liao, R. Teo, J. L. Wang, X. Dong, F. Lin, and K. Peng, "Distributed formation and reconfiguration control of VTOL UAVs," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 1, pp. 270–277, Jan. 2017.
- [131] K. Zhang, B. Jiang, and P. Shi, "Adjustable parameter-based distributed fault estimation observer design for multiagent systems with directed graphs," *IEEE Trans. Cybern.*, vol. 47, no. 2, pp. 306–314, Feb. 2017.
- [132] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia, PA, USA: SIAM, 2004.
- [133] Y. Hua, X. Dong, G. Hu, Q. Li, and Z. Ren, "Distributed time-varying output formation tracking for heterogeneous linear multiagent systems with a nonautonomous leader of unknown input," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4292–4299, Oct. 2019.
- [134] S. Zuo, Y. Song, F. L. Lewis, and A. Davoudi, "Optimal robust output containment of unknown heterogeneous multiagent system using off-policy reinforcement learning," *IEEE Trans. Cybern.*, vol. 48, no. 11, pp. 3197–3207, Nov. 2018.

- [135] J. Scherer *et al.*, "An autonomous multi-UAV system for search and rescue," in *Proc. 1st Workshop Micro Aerial Veh. Netw. Syst. Appl. Civil. Use*, 2015, pp. 33–38.
- [136] I. Maza, F. Caballero, J. Capitán, J. R. Martínez-de Dios, and A. Ollero, "Experimental results in multi-UAV coordination for disaster management and civil security applications," *J. Intell. Robot. Syst.*, vol. 61, nos. 1–4, pp. 563–585, 2011.
- [137] Y. Liu *et al.*, "A distributed control approach to formation balancing and maneuvering of multiple multirotor UAVs," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 870–882, Aug. 2018.
- [138] X. Dong, B. Yu, Z. Shi, and Y. Zhong, "Time-varying formation control for unmanned aerial vehicles: Theories and applications," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 1, pp. 340–348, Jan. 2015.
- [139] D. Xiwang, L. Yangfan, L. Chuang, H. Guoqiang, L. Qingdong, and Z. Ren, "Time-varying formation tracking for UAV swarm systems with switching directed topologies," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3674–3685, Dec. 2019.
- [140] G. Yang, S. Wang, J. Yang, B. Shen, and P. Shi, "Pose estimation of daily containers for a life-support robot," *Int. J. Innovat. Comput. Inf. Control*, vol. 14, no. 4, pp. 1545–1552, 2018.
- [141] J. Shao, G. Xie, and L. Wang, "Leader-following formation control of multiple mobile vehicles," *IET Control Theory Appl.*, vol. 1, no. 2, pp. 545–552, Mar. 2007.
- [142] X. Liang, H. Wang, Y. H. Liu, W. Chen, and T. Liu, "Formation control of nonholonomic mobile robots without position and velocity measurements," *IEEE Trans. Robot.*, vol. 34, no. 2, pp. 434–446, Apr. 2018.
- [143] Z. Li, Y. Yuan, F. Ke, W. He, and C.-Y. Su, "Robust vision-based tube model predictive control of multiple mobile robots for leader–follower formation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 3096–3106, Apr. 2020.
- [144] S. M. Lee, H. Kim, H. Myung, and X. Yao, "Cooperative coevolutionary algorithm-based model predictive control guaranteeing stability of multirobot formation," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 1, pp. 37–51, Jan. 2015.
- [145] X. Li and D. Zhu, "An adaptive SOM neural network method for distributed formation control of a group of AUVs," *IEEE Trans. Ind. Electron.*, vol. 65, no. 10, pp. 8260–8270, Oct. 2018.
- [146] C. Suryendu and B. Subudhi, "Modified constrained adaptive formation control scheme for autonomous underwater vehicles under communication delays," *IET Cyber Syst. Robot.*, vol. 2, no. 1, pp. 22–30, Mar. 2020.
- [147] X. Zhang, A. J. Flueck, and C. P. Nguyen, "Agent-based distributed volt/var control with distributed power flow solver in smart grid," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 600–607, Mar. 2016.
- [148] V. P. Singh, N. Kishor, and P. Samuel, "Distributed multi-agent system-based load frequency control for multi-area power system in smart grid," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5151–5160, Jun. 2017.
- [149] P. Wang and M. Govindarasu, "Multi-agent based attack-resilient system integrity protection for smart grid," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3447–3456, Jul. 2020.
- [150] B. U. I. Khan, F. Anwar, R. F. Olanrewaju, B. R. Pampori, and R. N. Mir, "A novel multi-agent and multilayered game formulation for intrusion detection in Internet of Things (IoT)," *IEEE Access*, vol. 8, pp. 98481–98490, 2020.
- [151] J. Wang, Y. Zhang, Y. Liu, and N. Wu, "Multiagent and bargaining-game-based real-time scheduling for Internet of Things-enabled flexible job shop," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2518–2531, Apr. 2019.
- [152] S. Lin, B. D. Schutter, Z. Zhou, and Y. Xi, "Multi-agent model-based predictive control for large-scale urban traffic networks using a serial scheme," *IET Control Theory Appl.*, vol. 9, no. 3, pp. 475–484, Feb. 2015.
- [153] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): Methodology and large-scale application on downtown Toronto," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1140–1150, Sep. 2013.
- [154] T. Wu *et al.*, "Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8243–8256, Aug. 2020.
- [155] C. Yu *et al.*, "Distributed multiagent coordinated learning for autonomous driving in highways based on dynamic coordination graphs," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 2, pp. 735–748, Feb. 2020.



Peng Shi (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Newcastle, Callaghan, NSW, Australia, in 1994, the Doctor of Science degree from the University of Glamorgan, Wales, U.K., in 2006, and the Doctor of Engineering degree from the University of Adelaide, Adelaide, SA, Australia, in 2015.

He is currently a Professor with the University of Adelaide. His research interests include automation and control systems, autonomous and robotic systems, cyber-physical systems, and network systems.

Prof. Shi has served on the Editorial Board for a number of journals, including *Automatica*, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, *IEEE TRANSACTIONS ON CYBERNETICS*, *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, and *IEEE CONTROL SYSTEMS LETTERS*. He is a Member of Board of Governors for IEEE SMC Society, and an IEEE Distinguished Lecturer. He is a Fellow of the Institution of Engineering and Technology, and the Institute of Engineers, Australia.



Bing Yan received the B.Sc. degree in automation and the M.Sc. degree in control theory and control engineering from the School of Automation, Northwestern Polytechnical University, Xi'an, China, in 2012 and 2015, respectively. She is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA, Australia.

Her research interests include flight control, formation control, and multiagent systems.

2.3 Supplementary literature review

2.3.1 Learning-based formation control

With the development of machine learning and computing technologies, learning-based formation control approaches have attracted the attention of many researchers in recent years. In practice, the system model is usually uncertain or even completely unknown, and traditional model-based or partial model-based control methods [45] are not adequate to solve control problems of complex unknown MAS. Learning-based control, also known as data-driven control, offers possible solutions to deal with unknown dynamics by designing the controller based entirely on experimental data collected from the plant [46]. Existing data-driven control methods includes model-free adaptive control (MFAC) methods [47–49], reinforcement learning (RL)-based control approaches [50,51], etc.

- MFAC methods have attracted significant attention in the field of MAS because they provide control systems with the ability to automatically learn and improve from experience [47–49]. A data-driven distributed formation control algorithm has been proposed for an unknown heterogeneous MAS based on MFAC to transform the unknown MAS into an equivalent virtual dynamic linearization data model [47]. The work in [49] developed an MFAC protocol for MAS to achieve formations and construct the agent models dynamically based on a linearization method only from input and output data. However, MFAC methods have limitations in the estimation of complex unknown dynamics when uncertainty or nonlinearity is too strong.
- RL has been proved to effectively achieve online optimization and remove complex nonlinear model information as it can maximize agents' reward in unknown environments [50, 51]. For instance, off-policy RL has been developed to deal with formation control problems [51–53] and output regulation problems [54,55]. Off-policy RL means the target policy differs from the behavior policy in RL, otherwise, it is called on-policy RL [56]. Compared with on-policy RL, off-policy RL has a great advantage in learning efficiency due to its parallel learning strategy. In recent years, off-policy deep RL algorithms are used to obtain the optimal control decision for high-order complex systems [57]. It combines the advantages of deep learning and reinforcement learning and directly allows the agent

to make decisions through the deep neural networks without the need to manually design the state space [58]. However, complex deep RL algorithms, such as deep deterministic policy gradient [59] and deep-Q-network [3], usually require high-performance computers to train a large amount of data offline. The issue on how to achieve real-time online performance when applying RL to practical distributed MAS with limited computing capability and energy demands further studies.

2.3.2 Collision-free formation optimization

Since the energy of agents is limited and collisions may occur in dynamic environments, optimal control with collision avoidance is another requirement that must be considered in formation control. Considerable advancements have been made for dynamic path planning using such approaches as heuristic [60] and artificial potential field (APF)-based approaches [61–63]. APF-based approaches have been widely used in multi-vehicle formations. In recent years, a control barrier function (CBF)-based safety-critical control has been proposed to protect systems from accident risks, such as avoiding collisions of autonomous vehicles [64]. In fact, the APF function has been proved to be a special case of CBF [65]. Two designs are provided to solve safety-critical problems. One is based on control Lyapunov function-control barrier function-quadratic programming (CLF-CBF-QP) to guarantee stability and safety simultaneously. The other is adding a safety filter based on control barrier function-quadratic programming (CBF-QP) after nominal controllers [66]. However, APF-based and CBF-based methods are still partially model-based, and only quadratic objective functions are considered in most existing optimization for collision avoidance. Therefore, how to design an optimal collision-free scheme for fully unknown MAS under non-quadratic objectives deserves further investigation.

Chapter 3

Robust Formation Control for Multi-agent Systems Based on Adaptive Observers

Statement of Authorship

Title of Paper	Robust formation control for multiagent systems based on adaptive observers
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	B. Yan, P. Shi, C. -C. Lim and C. Wu, "Robust formation control for multiagent systems based on adaptive observers," IEEE Systems Journal, vol. 16, no. 2, pp. 3139–3150, 2022.

Principal Author


Name of Principal Author (Candidate)	Bing Yan		
Contribution to the Paper	Conceptualization, methodology, validation and writing-original draft		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25 Jul 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Peng Shi		
Contribution to the Paper	Polishing, checking and verification		
Signature		Date	25 Jul 2022

Name of Co-Author	Cheng-Chew Lim		
Contribution to the Paper	Review, refine and validate		
Signature		Date	25 Jul 2022

Name of Co-Author	Chengfu Wu		
Contribution to the Paper	Review and editing		
Signature		Date	26 Jul 2022

3.1 Introduction

This chapter provides a novel distributed robust control strategy for uncertain heterogeneous multi-agent systems (MAS) to achieve time-varying formations (TVF) under switching topologies and multiple disturbances. Compared with mixed-order heterogeneous MAS (e.g. a MAS composed of first-order integrators and second-order integrators), a unified linear heterogeneous MAS with different orders and dynamics is considered in this chapter. An adaptive observer is developed under switching topologies to estimate the state information of a reference exosystem only based on local information, which is used for decoupling the heterogeneous dynamics from networks. Considering the physical system layer constraints of uncertainties, homogeneous disturbances, and heterogeneous disturbances, a robust L_2 controller is designed for unified heterogeneous MAS to achieve TVF. Finally, a case study of a UGV-UAV TVF for bushfire edge tracking and patrolling is presented. Comparative simulation results demonstrate that our solution has significant advantages in the case of the MAS against multiple disturbances.

3.2 Publication

B. Yan, P. Shi, C. -C. Lim and C. Wu, "Robust formation control for multiagent systems based on adaptive observers," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3139–3150, 2022.

Robust Formation Control for Multiagent Systems Based on Adaptive Observers

Bing Yan ¹, Peng Shi ¹, *Fellow, IEEE*, Cheng-Chew Lim ¹, *Life Senior Member, IEEE*, and Chengfu Wu

Abstract—In this article, a distributed robust control strategy is proposed for heterogeneous multiagent systems subject to uncertainties and disturbances to achieve time-varying formations under static and switching topologies. Without prior knowledge of global information on communication graphs, a distributed observer is developed to estimate the state information of a reference system. Internal model-based robust formation controllers for homogeneous disturbances, with and without unknown heterogeneous disturbances, are designed using the Riccati equation approach. Finally, a case study of bushfire edge tracking and patrolling is presented, and the effectiveness and robustness of the formation control developed strategy are verified by simulations.

Index Terms—Adaptive observers, heterogeneous multiagent systems, robust formation control, uncertainties and disturbances.

I. INTRODUCTION

RECENT decades have witnessed the rapid development of formation control for multiagent systems (MASs) in a wide range of their potential applications such as environmental monitoring, and disaster relief [1], [2]. Controlling formation can be fixed formation [3] or time-varying formation (TVF) [4], [5] depending on whether the reference formations change over time or not. Having the ability to vary formation references offers not only operational flexibility but also robustness in formation. There has been significant progress made in TVF modeling [5] and TVF controller designs [6] for MAS systems composed of agents with the same dynamics. However, modeling TVF to adapt to practical tasks and designing TVF controllers for MASs with different types of agents remain to be fully addressed.

Heterogeneous MASs composed of agents with different dynamics have advantages in terms of formation flexibility and complex task decomposition because of the different functions of individuals. Some typical heterogeneous systems were intensively studied, such as the mixed-order MASs [7]–[9], and the same-order MASs with different dynamics [10]. In recent years, the consensus problems on unified heterogeneous MASs with different orders and dynamics has drawn much attention [11]. The output regulation control was first introduced to solve output

consensus problems for heterogeneous MASs by following a common exosystem. The exosystem refers to an external generator of reference (and/or disturbance) signals for the output of the plant to track (and/or reject) [12]. The output regulation control theory has led to the investigation of the output consensus problems [13] and formation control [14] for heterogeneous MASs.

The approaches used for formation control are mostly centralized control [15] and distributed control [16], [17]. The former requires a control center and global information. It lacks robustness against a center failure and scales poorly. Distributed formation control based on local information is more flexible in scalability and better fault-tolerant toward partial failures. There are a number of formation control methods that can be broadly classified as distributed based on adaptive observers to estimate the global information of the reference system [18]–[21]. However, some existing observer designs still require global topology information to estimate the states or outputs of the reference exosystems [21], [22]. Meanwhile, the communication topologies are also required to have properties that enable the use of a static graph for node-based adaptive observers [23], or connected graphs for edge-based adaptive observers to track marginally stable exosystems under switching topologies [24]. In particular, for large-scale systems under switching topologies, global information of topologies is generally difficult to obtain, and the topology properties required for the observer design are also hard to guarantee in practical applications. Therefore, more intensive research is requested on the observer design for the distributed TVF control of heterogeneous MASs under unavailable global communication information and switching topologies, especially when tracking unstable exosystems.

Achieving robustness against uncertainties and disturbances in formation control is important for MASs in real-world applications. Considerable efforts based on robust control have been made for heterogeneous MASs [12], [23], [25], [26]. Since disturbances are regarded as part of the exosystem in the output regulation framework [12], existing methods in [23] and [26] for heterogeneous MASs are only applicable to the situation in which each agent is subject to the same disturbance. We call it homogeneous disturbance for all agents in a common environment. In addition to it, the local disturbances on each agent are generally heterogeneous, unknown and unmodeled, and we call them heterogeneous disturbances. Using the bushfire tracking and patrolling task as an example, the direction of the main wind in the fire area will affect the movements of the whole MASs, which can be regarded as a homogeneous disturbance. Meanwhile, disturbances of heterogeneity and model uncertainties occur due to different fluctuations of local temperature and air pressure. As the model information of exosystems is required to be known at least by a subset of agents [18],

Manuscript received December 9, 2020; revised May 27, 2021 and October 1, 2021; accepted November 7, 2021. Date of publication December 2, 2021; date of current version June 13, 2022. This work was supported by Australian Research Council under Grant DP170102644. (*Corresponding author: Peng Shi.*)

Bing Yan, Peng Shi, and Cheng-Chew Lim are with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia (e-mail: bing.yan@adelaide.edu.au; peng.shi@adelaide.edu.au; cheng.lim@adelaide.edu.au).

Chengfu Wu is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: wu.nwpu@gamil.com).

Digital Object Identifier 10.1109/JSYST.2021.3127579

unknown and unmodeled disturbances cannot be observed as part of exosystems, and there are not as trivial to handle them in the output regulation framework. To the best of our knowledge, the robust formation control problem for heterogeneous MASs under multiple disturbances has not been fully investigated, which motivates us for the current article.

Overall, this investigation has the following main contributions.

1) Compared with mixed-order heterogeneous MASs [7], [8], a TVF control strategy is presented for a more general linear heterogeneous MAS with different orders and dynamics to adapt to complex and realistic tasks such as those found in bushfire tracking and patrolling.

2) Compared with the research conducted in [7] and [22], a distributed adaptive observer is designed to estimate the states of the reference system without requiring the global information of the communication topology, so that the dynamics of n agents can be decoupled from the network. The design is expanded and is applicable to any switching topologies under the spanning tree assumption (Assumption 2) with an average dwell time-based condition.

3) Differing from [23] and [26], both homogeneous disturbance and heterogeneous disturbances are considered in the proposed bounded L_2 controller for heterogeneous uncertain MASs. Comparative simulations demonstrate that our solution can effectively improve the robustness under unknown heterogeneous disturbances.

The notation used is standard throughout this article. \otimes stands for the Kronecker product. $\|X\|$ is the norm of a matrix X or a vector X . Matrices $\mathbf{0}$ and \mathbf{I} are all-zeros matrix and unit matrix with corresponding dimensions. $\text{diag}(v)$ represents a diagonal matrix constructed by the vector v . $\text{blkdiag}(A_1, A_2, \dots, A_n)$ denotes the block diagonal matrix created by aligning the input matrices. $\text{col}(x_1, x_2, \dots, x_n)$ is the column vector consisting of vector x_1 to vector x_n .

II. PROBLEM FORMULATION

A heterogeneous MAS with n agents of different orders and dynamics can be modeled as

$$\dot{x}_i = A_i x_i + B_i u_i, \quad y_i = C_i x_i, \quad i = 1, 2, \dots, n \quad (1)$$

where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}$, and $y_i \in \mathbb{R}^{p_i}$ represent the state, input, and output variables of the i th agent, respectively. Matrices A_i , B_i , and C_i are the system state, input, and output matrices, respectively. When system uncertainties and disturbances are taking into account, system (1) has the following form:

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{D}_i \omega + E_i \varpi_i \\ y_i &= C_i x_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (2)$$

where matrices

$$\tilde{A}_i = A_i + \Delta A_i, \quad \tilde{B}_i = B_i + \Delta B_i, \quad \tilde{D}_i = D_i + \Delta D_i$$

refer to the system matrices $\tilde{\star}_i$, consisting of the normal matrices \star_i and uncertainty matrices $\Delta \star_i$ of the i th agent, where $\star = A, B, D$. The uncertainty matrices are assumed to be bounded, and belong to an open neighborhood W of the original point. Note that the set of W does not need to be small. Similar assumptions have been found in [12] and [25].

The disturbance considered here is homogeneous disturbance $\omega \in \mathbb{R}^q$ and heterogeneous disturbance $\varpi_i \in \mathbb{R}^{q_i}$. When the

dynamics of the disturbance like environment disturbance acts identically on n agents, we call it the homogeneous disturbance which generated by

$$\dot{\omega} = A_\omega \omega \quad (3)$$

where the state matrix A_ω is assumed to be known as its dynamics can be approximated by a Fourier series. The input matrix of ω in (2) is \tilde{D}_i .

When the disturbance acting on n agents is different, it is called the heterogeneous disturbance ϖ_i . The input matrix of the heterogeneous disturbance is E_i . For some highly nonlinear unknown disturbances, we regard them as heterogeneous disturbance ϖ_i which is assumed to be unknown and unobservable but bounded.

In the desired formation, a virtual leader which acts as the root node of the formation graph is modeled as

$$\dot{x}_0 = A_0 x_0, \quad y_0 = C_0 x_0 \quad (4)$$

where $x_0 \in \mathbb{R}^{n_0}$ and $y_0 \in \mathbb{R}^{p_0}$ denote the system state and output variables, respectively, with A_0 and C_0 being the system state and output matrices, respectively. Note that there is no external control input to the virtual leader, and it is assumed that only a subset of agents can get access to the virtual leader.

To define the shape of the formation for n agents, it requires n TVF systems to output the time-varying displacement of each agent relative to the virtual leader. The n TVF systems are designed as

$$\dot{f}_i = A_i^f f_i, \quad y_i^f = C_i^f f_i, \quad i = 1, 2, \dots, n \quad (5)$$

where $f_i \in \mathbb{R}^{n_f}$ and $y_i^f \in \mathbb{R}^{p_f}$ are the state and output variables of the TVF, respectively. Matrices A_i^f and C_i^f are the state and output matrices, respectively. Note that each TVF is only designed for each agent, it is assumed that every agent knows its respective TVF.

Remark 1: Although the agents have different orders and different dynamics, they can still send their variables of common interest such as positions as their outputs. Since it is difficult to ensure that the states with different orders achieve consensus, it is more practical to investigate, for the heterogeneous MASs, the output consensus problems than the state consensus problems. Therefore, we assume that the dimension of the position outputs $p_i = p_0 = p_f$ of systems (1), (2), (4), and (5) is the same. The assumption is pertinent to formation control problems of MASs when completing different tasks. An example of a tracking and patrolling task is given in the simulation section to illustrate the design of the virtual leader and the TVF.

We introduce the following definitions, control problems and assumptions for developing the main result given in Section III.

Definition 1: For any initial condition, system (1) with a virtual leader (4) is said to achieve output consensus if the following condition holds:

$$\lim_{t \rightarrow \infty} \|C_i x_i - C_0 x_0\| = \mathbf{0}, \quad i = 1, 2, \dots, n. \quad (6)$$

Definition 2: For any initial condition, system (1) with a virtual leader (4) and a reference formation (5) is said to achieve formation output consensus if the following condition holds:

$$\lim_{t \rightarrow \infty} \|C_i x_i - C_0 x_0 - C_i^f f_i\| = \mathbf{0}, \quad i = 1, 2, \dots, n. \quad (7)$$

Problem 1: Robust formation output control problem is to design a controller to meet the following condition:

$$\lim_{t \rightarrow \infty} \|e_i\| = \mathbf{0}, \quad i = 1, 2, \dots, n \quad (8)$$

where $e_i = C_i x_i - C_0 x_0 - C_i^f f_i$ is the formation error, for system (2) with a virtual leader (4) and a reference formation (5), when $\varpi_i = 0$, and $\Delta_{\star_i} \in W$, where W is an open neighborhood of $\Delta_{\star_i} = \mathbf{0}$.

Problem 2: Bounded L_2 robust formation output control problem is to design a controller to meet the following condition:

$$\int_0^\infty e_i^T e_i dt \leq \gamma_i^2 \int_0^\infty \varpi_i^T \varpi_i dt, \quad i = 1, 2, \dots, n \quad (9)$$

for system (2) with virtual leader (4) and reference formation (5), when $\varpi_i \in L_2[0, +\infty)$ [27], and $\Delta_{\star_i} \in W$, where W is an open neighborhood of $\Delta_{\star_i} = \mathbf{0}$.

The information flow among the MAS composed of n agents and the virtual leader is generally described based on graph theory [28]. We consider two cases: (a) The graph of the MAS G_s is a static direct graph G ; and (b) G_s is a dynamic direct graph $G^{\sigma(t)}$ that is switched among $\{G^1, G^2, \dots, G^M\}$ under a switching signal $\sigma(t) : [0, +\infty) \rightarrow \{1, 2, \dots, M\}$ [29].

Assumption 1: Matrix $A_i^v = \text{blkdiag}(A_0, A_\omega, A_i^f)$ has no eigenvalues with negative real part.

Assumption 2: (a) [21] If $G_s = G$, assume that graph G has a spanning tree and the virtual leader is the root node.

(b) [29] If $G_s = G^{\sigma(t)}$, assume that there exists an infinite sequence of uniformly bounded, nonoverlapping time intervals $[t_k, t_{k+1})$ with $t_{k+1} - t_k < g$ for some positive g , over which the graph is time-invariant. Each graph G^i , $i \in \{1, 2, \dots, M\}$ has a spanning tree and the virtual leader is the root node.

Assumption 3: The pair (A_i, B_i) is stabilizable.

Assumption 4: For any $\lambda \in \sigma(A_i^v)$,

$$\text{rank} \begin{bmatrix} A_i - \lambda I_{n_i} & B_i \\ C_i & \mathbf{0} \end{bmatrix} = n_i + p_i, \quad i = 1, 2, \dots, n$$

where $\sigma(A_i^v)$ denotes the spectrum of A_i^v .

Remark 2: Note that Assumption 1 is to avoid collisions among agents caused by tracking stable virtual leader and TVF to zero, and ensure no disturbances will automatically disappear. We can see similar assumptions in [23]. Different requirements for information consensus under static and switching topologies are given in Assumption 2 [21], [29]. Note that the existence of the common internal model dynamics of heterogeneous agents is guaranteed by Assumptions 4 [12], [30].

III. MAIN RESULTS

In this section, a distributed adaptive observer is proposed for a common reference exosystem to decouple the dynamics of the agents from static and switching communication topologies. Then, two robust formation control strategies are designed to solve Problems 1 and 2, respectively.

A. Distributed Adaptive Observer

For the robust formation output control problems, a common reference exosystem consists of the virtual leader and the homogeneous disturbance as

$$\dot{\xi} = A^\xi \xi, \quad y^\xi = C^\xi \xi \quad (10)$$

where the state variable is $\xi = \text{col}(x_0, \omega)$, and $A^\xi = \text{blkdiag}(A_0, A_\omega)$ and $C^\xi = [C_0, \mathbf{0}]$ are the system state and output matrices, respectively. The formation error is rewritten as

$$e_i = C_i x_i - C^\xi \xi - C_i^f f_i. \quad (11)$$

By Assumption 2, we index the virtual leader by number 0, and the Laplacian matrices of G and $G^{\sigma(t)}$ are

$$L_{n+1} = \begin{bmatrix} 0 & \mathbf{0}_{1 \times n} \\ L_0 & L_1 \end{bmatrix}, \quad L_{n+1}^{\sigma(t)} = \begin{bmatrix} 0 & \mathbf{0}_{1 \times n} \\ L_0^{\sigma(t)} & L_1^{\sigma(t)} \end{bmatrix} \quad (12)$$

where $L_1 = L + A_{n_0}$, and $L_1^{\sigma(t)} = L^{\sigma(t)} + A_{n_0}^{\sigma(t)}$. The Laplacian matrices only for n agents under G and $G^{\sigma(t)}$ are L and $L^{\sigma(t)}$, respectively. The Laplacian matrices between the virtual leader and n agents under G and $G^{\sigma(t)}$ are L_0 and $L_0^{\sigma(t)}$, respectively. The adjacency matrices corresponding to Laplacian matrices (L, L_0) and $(L^{\sigma(t)}, L_0^{\sigma(t)})$ are $(A_{nn} = [a_{ij}], A_{n0} = [a_{i0}])$ and $(A_{nn}^{\sigma(t)} = [a_{ij}(t)], A_{n0}^{\sigma(t)} = [a_{i0}(t)])$, respectively.

We design a distributed adaptive observer to estimate states of virtual leader and homogeneous disturbance as

$$\begin{aligned} \dot{\hat{\xi}}_i &= A^\xi \hat{\xi}_i - (R^{-1}P)(\theta_i + \phi_i)\psi_i, \quad i = 1, 2, \dots, n \\ \psi_i &= \sum_{j=1}^n a_{ij} (\hat{\xi}_i - \hat{\xi}_j) + a_{i0} (\hat{\xi}_i - \xi), \quad G_s = G \\ \psi_i &= \sum_{j=1}^n a_{ij}(t) (\hat{\xi}_i - \hat{\xi}_j) + a_{i0}(t) (\hat{\xi}_i - \xi), \quad G_s = G^{\sigma(t)} \\ \dot{\theta}_i &= \psi_i^T (PR^{-1}P)\psi_i \quad \theta_i(0) > 0 \\ \phi_i &= \psi_i^T P\psi_i \end{aligned} \quad (13)$$

where $\hat{\xi}_i$ denotes the estimation of ξ for i th agent. Matrix P is a positive definite matrix that satisfies the inequality

$$(A^\xi)^T P + P A^\xi - P R^{-1} P < 0 \quad (14)$$

and R is a given symmetric and positive definite matrix.

We define the observation error as $\tilde{\xi}_i = \hat{\xi}_i - \xi$ for i th agent, and denote the related vectors and matrices of the MAS by

$$\begin{aligned} \underline{\xi} &= \mathbf{I}_n \otimes \xi, \quad \hat{\underline{\xi}} = \text{col}(\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_n), \quad \tilde{\underline{\xi}} = \text{col}(\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_n) \\ \Theta &= \text{diag}(\theta_1, \theta_2, \dots, \theta_n), \quad \Phi = \text{diag}(\phi_1, \phi_2, \dots, \phi_n) \\ \Psi &= \text{col}(\psi_1, \psi_2, \dots, \psi_n), \quad \varphi = \text{col}(\phi_1, \phi_2, \dots, \phi_n). \end{aligned}$$

It follows from (12) that

$$\begin{aligned} \Psi &= (L_1 \otimes \mathbf{I}) \hat{\underline{\xi}} + (L_0 \otimes \mathbf{I}) \underline{\xi} = (L_1 \otimes \mathbf{I}) \tilde{\underline{\xi}}, \quad G_s = G \\ \Psi &= (L_1^{\sigma(t)} \otimes \mathbf{I}) \hat{\underline{\xi}} + (L_0^{\sigma(t)} \otimes \mathbf{I}) \underline{\xi} = (L_1^{\sigma(t)} \otimes \mathbf{I}) \tilde{\underline{\xi}}, \quad G_s = G^{\sigma(t)}. \end{aligned} \quad (15)$$

Based on (13), the derivatives of variables $\tilde{\xi}$, Ψ , and φ under static and switching topologies are

$$\begin{cases} \dot{\tilde{\xi}} = (\mathbf{I}_n \otimes A^\xi - (\Theta + \Phi)L_1 \otimes R^{-1}P)\tilde{\xi} \\ \dot{\Psi} = (\mathbf{I}_n \otimes A^\xi - (\Theta + \Phi)L_1 \otimes R^{-1}P)\Psi \\ \quad = (L_1 \otimes \mathbf{I})\dot{\tilde{\xi}}, \quad G_s = G \\ \dot{\varphi} = \Psi^T [\mathbf{I}_n \otimes (PA^\xi + (A^\xi)^T P) \\ \quad - (L_1^T + L_1)(\Theta + \Phi) \otimes PR^{-1}P] \Psi \end{cases}$$

$$\begin{cases} \dot{\tilde{\xi}} = (\mathbf{I}_n \otimes A^\xi - (\Theta + \Phi)L_1^{\sigma(t)} \otimes R^{-1}P)\tilde{\xi}, \quad G_s = G^{\sigma(t)} \\ \dot{\Psi} = (\mathbf{I}_n \otimes A^\xi - (\Theta + \Phi)L_1^{\sigma(t)} \otimes R^{-1}P)\Psi \\ \quad = (L_1^{\sigma(t)} \otimes \mathbf{I})\dot{\tilde{\xi}}, \quad G_s = G^{\sigma(t)}, t \in [t_k, t_{k+1}) \\ \dot{\varphi} = \Psi^T \left[- \left((L_1^{\sigma(t)})^T + L_1^{\sigma(t)} \right) (\Theta + \Phi) \otimes PR^{-1}P \right. \\ \quad \left. + \mathbf{I}_n \otimes (PA^\xi + (A^\xi)^T P) \right] \Psi, G_s = G^{\sigma(t)}, t \in [t_k, t_{k+1}) \end{cases} \quad (16)$$

Remark 3: Note that observer (13) can still be applied in the absence of disturbance by setting $\omega = 0$. The matrix $C^\xi = [C_0, \mathbf{0}]$ in (10) is designed to ensure that the error signal (11) used for formation control is not affected by disturbance. Therefore, the observer is suitable for controller design with or without disturbance.

We recall the following results in order to derive our main results in the sequel.

Lemma 1 [31]: Under Assumption 2, all the eigenvalues of L_1 and $L_1^{\sigma(t)}$ have positive real parts, and there exist symmetric positive definite diagonal matrices $M = \text{diag}(m_1, m_2, \dots, m_n)$, $m_i > 0$ and $M^{\sigma(t)} = \text{diag}(m_1^{\sigma(t)}, m_2^{\sigma(t)}, \dots, m_n^{\sigma(t)})$, $m_i^{\sigma(t)} > 0$ satisfying $ML_1 + L_1^T M \geq \lambda_0 \mathbf{I}$ and $M^{\sigma(t)} L_1^{\sigma(t)} + (L_1^{\sigma(t)})^T M^{\sigma(t)} \geq \lambda_0^{\sigma(t)} \mathbf{I}$, where $\lambda_0 > 0$ and $\lambda_0^{\sigma(t)} > 0$ are the minimum eigenvalues of $ML_1 + L_1^T M$ and $M^{\sigma(t)} L_1^{\sigma(t)} + (L_1^{\sigma(t)})^T M^{\sigma(t)}$, respectively.

Lemma 2 [32] (Young's inequality): Assuming that a, b, p , and q are positive real numbers, and p and q are such that $\frac{1}{p} + \frac{1}{q} = 1$, then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.

Lemma 3 [12]: The pair (Σ_1, Σ_2) incorporates the p -copy internal model of any given square matrix A with any given integer $p > 0$, if the pair satisfy the form

$$\Sigma_1 = T \begin{bmatrix} S_1 & S_2 \\ 0 & \bar{\Sigma}_1 \end{bmatrix} T^{-1}, \quad \Sigma_2 = T \begin{bmatrix} S_3 \\ \bar{\Sigma}_2 \end{bmatrix} \quad (17)$$

where T is any nonsingular matrix, S_1, S_2 , and S_3 are any constant matrices,

$\bar{\Sigma}_1 = \text{blkdiag}(\beta_{11}, \dots, \beta_{1p})$ and $\bar{\Sigma}_2 = \text{blkdiag}(\beta_{21}, \dots, \beta_{2p})$.

If the minimal polynomial of A is expressed as

$$\lambda^n + \alpha_1 \lambda^{(n-1)} + \dots + \alpha_{(n-1)} \lambda + \alpha_n$$

then (β_{1j}, β_{2j}) is stabilizable for $j = 1, 2, \dots, p$ in the form of

$$\beta_{1j} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -\alpha_n & -\alpha_{(n-1)} & \dots & -\alpha_1 \end{bmatrix}, \beta_{2j} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \quad (18)$$

B. Robust Formation Output Control Problems

Based on the distributed adaptive observer (13), the robust formation output controllers are designed under system uncertainties and homogeneous disturbance, with and without subjecting to additional heterogeneous disturbance.

When heterogeneous disturbance $\varpi_i = 0$, the heterogeneous MAS (2) is simplified to

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{D}_i \omega \\ y_i &= C_i x_i, \quad i = 1, 2, \dots, n. \end{aligned} \quad (19)$$

Considering the n TVF in (5), the n extended exosystems are defined as

$$\dot{v}_i = A_i^v v_i, \quad y_i^v = C_i^v v_i \quad (20)$$

where the system state variable is $v_i = \text{col}(\xi, f_i)$, with $A_i^v = \text{blkdiag}(A^\xi, A_i^f)$, and $C_i^v = [C^\xi, C_i^f]$ representing the system state and output matrices, respectively.

The augmented system of (19), (20), and (11) is

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{F}_i v_i \\ y_i &= C_i x_i, \quad i = 1, 2, \dots, n \\ e_i &= C_i x_i - S_i v_i \end{aligned} \quad (21)$$

where

$$\tilde{F}_i = [\mathbf{0}, \tilde{D}_i, \mathbf{0}], \quad S_i = [C_0, \mathbf{0}, C_i^f]$$

The state feedback control law is designed as

$$\begin{aligned} u_i &= K_i^x x_i + K_i^z z_i \\ \dot{z}_i &= \Sigma_i^1 z_i + \Sigma_i^2 (C_i x_i - C^\xi \hat{\xi}_i - C_i^f f_i) \end{aligned} \quad (22)$$

where z_i is the state of the dynamic compensator, K_i^x and K_i^z are the control laws for x_i and z_i , and (Σ_i^1, Σ_i^2) is the p -copy internal model pair of A_i^v constructed by Lemma 3, based on the internal model principle [12]. Construct the Riccati equation as

$$(A_i^\delta)^T P_i + P_i A_i^\delta - \frac{1}{\epsilon_{i1}} P_i B_i^\delta R_i^{-1} (B_i^\delta)^T P_i + (\epsilon_{i1} + \epsilon_{i2}) Q_i = 0 \quad (23)$$

where

$$A_i^\delta = \begin{bmatrix} A_i & 0 \\ \Sigma_i^2 C_i & \Sigma_i^1 \end{bmatrix}, \quad B_i^\delta = \begin{bmatrix} B_i \\ \mathbf{0} \end{bmatrix}.$$

The control law is chosen as

$$K_i = [K_i^x \ K_i^z] = -\frac{1}{2\epsilon_{i1}} R_i^{-1} B_i^T P_i \quad (24)$$

where matrix P_i is the symmetric positive definite solution of the Riccati (23), for given $\epsilon_{i1} > 0$, $\epsilon_{i2} > 0$ and symmetric positive definite matrices Q_i and R_i .

In the presence of additional heterogeneous disturbance, the heterogeneous MAS is recast as

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{F}_i v_i + E_i \varpi_i \\ y_i &= C_i x_i, \quad i = 1, 2, \dots, n \\ e_i &= C_i x_i - S_i v_i. \end{aligned} \quad (25)$$

The structure of the robust control law is the same as that in (22)–(24), but matrix P_i is the solution of the associated Riccati equation

$$\begin{aligned} (A_i^\delta)^T P_i + P_i A_i^\delta - \frac{1}{\epsilon_{i1}} P_i B_i^\delta R_i^{-1} (B_i^\delta)^T P_i + \frac{1}{\gamma_i^2} P_i E_i E_i^T P_i \\ + \frac{1}{\gamma_i^2} C_i^T C_i + (\epsilon_{i1} + \epsilon_{i2}) Q_i = 0 \end{aligned} \quad (26)$$

for given $\epsilon_{i1} > 0$, $\epsilon_{i2} > 0$, $0 < \gamma_i < 1$ and symmetric positive definite matrices Q_i and R_i .

The main result of this article follows.

Theorem 1: Consider systems (2)–(5) satisfying Assumptions 1–4. (a) When $\varpi_i = 0$ and $G_s = G$, Problem 1 is solvable by distributed observer (13) and state feedback controller (22), if conditions (14) and (23) are satisfied.

(b) When $\varpi_i = 0$ and $G_s = G^{\sigma(t)}$, Problem 1 is solvable by distributed observer (13) and state feedback controller (22) under conditions (14) and (23), if the average dwell time of each agent is longer than a positive threshold that can be decreased by choosing a sufficiently large initial value of θ_i .

Proof: (a) For the situation that $\varpi_i = 0$ and $G_s = G$, substituting (24) into (23) gives

$$(A_i^c)^T P_i + P_i A_i^c < -(\epsilon_{i1} + \epsilon_{i2}) Q_i < 0 \quad (27)$$

where

$$A_i^c = A_i^\delta + B_i^\delta K_i = \begin{bmatrix} A_i + B_i K_i^x & B_i K_i^z \\ \Sigma_i^2 C_i & \Sigma_i^1 \end{bmatrix}.$$

Thus, A_i^c is Hurwitz from (27). From [12, Lemma 1.20], if A_i^c is Hurwitz, then \tilde{A}_i^c is also Hurwitz under Assumption 4, for any $\Delta_{\star_i} \in W$.

Substituting of state feedback law (22) into uncertain heterogeneous MAS (21) yields

$$\begin{aligned} \dot{\delta}_i &= \tilde{A}_i^c \delta_i + \tilde{B}_i^c v_i + J_i^\delta \tilde{\xi}_i \\ e_i &= C_i^c \delta_i - S_i v_i \end{aligned} \quad (28)$$

where $\delta_i = \text{col}(x_i, z_i)$, and

$$\tilde{B}_i^c = \begin{bmatrix} \tilde{F}_i \\ -\Sigma_i^2 C_i^v \end{bmatrix}, \quad J_i^\delta = \begin{bmatrix} \mathbf{0} \\ -\Sigma_i^2 C_i^\xi \end{bmatrix}, \quad C_i^c = [C_i \ \mathbf{0}].$$

Based on [12, Lemma 1.27], if A_i^c is Hurwitz and (Σ_i^1, Σ_i^2) incorporates a p -copy internal model of A_i^v , the following equations:

$$\begin{aligned} \Pi_i^x A_i^v &= (\tilde{A}_i + \tilde{B}_i K_i^x) \Pi_i^x + \tilde{B}_i K_i^z \Pi_i^z + \tilde{F}_i \\ \Pi_i^z A_i^v &= \Sigma_i^1 \Pi_i^z + \Sigma_i^2 (C_i \Pi_i^x - S_i) \end{aligned} \quad (29)$$

have a unique solution (Π_i^x, Π_i^z) , which satisfies

$$0 = C_i \Pi_i^x - S_i. \quad (30)$$

Constructing a new variable $\Pi_i^v = \text{col}(\Pi_i^x, \Pi_i^z)$ and substituting it into (29) and (30) leads to

$$\begin{aligned} \Pi_i^v A_i^v &= \tilde{A}_i^c \Pi_i^v + \tilde{B}_i^c \\ 0 &= C_i^c \Pi_i^v - S_i. \end{aligned} \quad (31)$$

From system (21) and (31), the dynamics of the state error $\tilde{\delta}_i = \delta_i - \Pi_i^v v_i$ is

$$\begin{aligned} \dot{\tilde{\delta}}_i &= \tilde{A}_i^c \tilde{\delta}_i + \tilde{B}_i^c v_i + J_i^\delta \tilde{\xi}_i - \Pi_i^v A_i^v v_i \\ &= \tilde{A}_i^c \tilde{\delta}_i + \tilde{B}_i^c v_i + J_i^\delta \tilde{\xi}_i - \tilde{A}_i^c \Pi_i^v v_i - \tilde{B}_i^c v_i \\ &= \tilde{A}_i^c \tilde{\delta}_i + J_i^\delta \tilde{\xi}_i. \end{aligned} \quad (32)$$

Similarly, the formation error in Problem 1 and Problem 2 is rewritten as

$$e_i = C_i^c \delta_i - S_i v_i = C_i^c \tilde{\delta}_i + (C_i^c \Pi_i^v - S_i) v_i = C_i^c \tilde{\delta}_i. \quad (33)$$

Under the controller in (22), closed-loop system (21) can be expressed as

$$\begin{aligned} \dot{\tilde{\delta}}_i &= (\tilde{A}_i^\delta + \tilde{B}_i^\delta K_i^x) \tilde{\delta}_i + J_i^\delta \tilde{\xi}_i = \tilde{A}_i^\delta \tilde{\delta}_i + J_i^\delta \tilde{\xi}_i \\ e_i &= C_i^c \tilde{\delta}_i, \quad i = 1, 2, \dots, n. \end{aligned} \quad (34)$$

We define Lyapunov functions as

$$V = V_1 + V_2 + V_3 \quad (35)$$

where

$$V_1 = \sum_{i=1}^n (\tilde{\delta}_i)^T P_i \tilde{\delta}_i \quad (36)$$

$$V_2 = \sum_{i=1}^n \frac{m_i}{2} (2\theta_i + \phi_i) \phi_i \quad (37)$$

$$V_3 = \sum_{i=1}^n \frac{m_i}{2} (\theta_i - a_1 - a_2)^2 \quad (38)$$

where m_i is defined in Lemma 1, and a_1 and a_2 are constants to be decided later.

From (13) and (22), it follows:

$$m_i > 0, \theta_i > 0, \phi_i \geq 0, P_i > 0$$

Therefore, $V \geq 0$.

Differentiating V_1 along the trajectory of (36) gives

$$\begin{aligned} \dot{V}_1 &= \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^\delta + \tilde{B}_i^\delta K_i^x)^T P_i + P_i (\tilde{A}_i^\delta + \tilde{B}_i^\delta K_i^x) \right] \tilde{\delta}_i \\ &\quad + \sum_{i=1}^n \tilde{\xi}_i^T (J_i^\delta)^T P_i \tilde{\delta}_i + \sum_{i=1}^n \tilde{\delta}_i^T P_i J_i^\delta \tilde{\xi}_i. \end{aligned} \quad (39)$$

Substituting feedback control law (24) into (39) results in

$$\begin{aligned} \dot{V}_1 &= \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^\delta)^T P_i + P_i \tilde{A}_i^\delta - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i \right. \\ &\quad \left. - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i \right] \tilde{\delta}_i \end{aligned}$$

$$+ \sum_{i=1}^n \tilde{\xi}_i^T (J_i^\delta)^T P_i \tilde{\delta}_i + \sum_{i=1}^n \tilde{\delta}_i^T P_i J_i^\delta \tilde{\xi}_i^T. \quad (40)$$

From Lemma 2, using Young's inequality obtains

$$\begin{aligned} \dot{V}_1 \leq & \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^\delta)^T P_i + P_i \tilde{A}_i^\delta - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i \right. \\ & \left. - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i + \epsilon_{i1} Q_i \right] \tilde{\delta}_i \\ & + \sum_{i=1}^n \tilde{\xi}_i^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|(J_i^\delta)^T J_i^\delta\| \right] \tilde{\xi}_i. \end{aligned} \quad (41)$$

Substituting (13) and (16) into the derivation of V_2 gives

$$\begin{aligned} \dot{V}_2 = & \sum_{i=1}^n [m_i(\theta_i + \phi_i)\dot{\phi}_i + m_i\dot{\theta}_i\phi_i] \\ = & \Psi^T [M(\Theta + \Phi) \otimes (PA^\xi + (A^\xi)^T P) \\ & - M(\Theta + \Phi)(L_1^T + L_1)(\Theta + \Phi) \otimes PR^{-1}P] \Psi \\ & + \Psi^T [M\Phi \otimes PR^{-1}P] \Psi. \end{aligned} \quad (42)$$

Under Assumption 2 and Lemma 1, applying inequality $L_1 M + M L_1 > \lambda_0 \mathbf{I}$ to (42) yields

$$\begin{aligned} \dot{V}_2 \leq & \Psi^T [M(\Theta + \Phi) \otimes (PA^\xi + (A^\xi)^T P) \\ & - \lambda_0(\Theta + \Phi)^2 \otimes PR^{-1}P] \Psi \\ & + \Psi^T [M\Phi \otimes PR^{-1}P] \Psi. \end{aligned} \quad (43)$$

Similarly, taking the derivation of V_3 along a trajectory in (38) implies

$$\begin{aligned} \dot{V}_3 = & \sum_{i=1}^n m_i(\theta_i - a_1 - a_2)\dot{\theta}_i \\ = & \Psi^T [(M\Theta - a_1 M - a_2 M) \otimes PR^{-1}P] \Psi \\ = & \Psi^T [(M\Theta - a_1 M) \otimes PR^{-1}P] \Psi \\ & - \tilde{\xi}^T [a_2 M L_1^T L_1 \otimes PR^{-1}P] \tilde{\xi}. \end{aligned} \quad (44)$$

Based on (41)–(44), we have

$$\begin{aligned} \dot{V} = & \dot{V}_1 + \dot{V}_2 + \dot{V}_3 \\ \leq & \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^\delta)^T P_i + P_i \tilde{A}_i^\delta - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i \right. \\ & \left. - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i + \epsilon_{i1} Q_i \right] \tilde{\delta}_i \\ & + \Psi^T [M(\Theta + \Phi) \otimes (PA^\xi + (A^\xi)^T P) \\ & - \lambda_0(\Theta + \Phi)^2 \otimes PR^{-1}P] \Psi \\ & + \Psi^T [(M\Theta + M\Phi - a_1 M) \otimes PR^{-1}P] \Psi \\ & + \tilde{\xi}^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|(J_i^\delta)^T J_i^\delta\| \right. \\ & \left. - a_2 M L_1^T L_1 \otimes PR^{-1}P \right] \tilde{\xi}. \end{aligned} \quad (45)$$

If condition (23) is met, then

$$\begin{aligned} (A_i^\delta)^T P_i + P_i A_i^\delta - \frac{1}{\epsilon_{i1}} P_i B_i^\delta R_i^{-1} (B_i^\delta)^T P_i \\ + \epsilon_{i1} Q_i = -\epsilon_{i2} Q_i < 0. \end{aligned} \quad (46)$$

From (46) and [12, Lemma 1.20], we can see that

$$\begin{aligned} (\tilde{A}_i^\delta)^T P_i + P_i \tilde{A}_i^\delta - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i \\ - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i + \epsilon_{i1} Q_i = -\epsilon_{i2} Q_i < 0. \end{aligned} \quad (47)$$

There exists a parameter a_1 to satisfy $a_1 > \frac{5m_i}{2\lambda_0}$, based on condition (14) and Young's inequality in Lemma 2, then we have

$$M(\Theta + \Phi) = \frac{M}{\sqrt{\lambda_0}} \sqrt{\lambda_0}(\Theta + \Phi) \leq \frac{M^2}{2\lambda_0} + \frac{\lambda_0(\Theta + \Phi)^2}{2}$$

and

$$\begin{aligned} & \Psi^T [M(\Theta + \Phi) \otimes (PA^\xi + (A^\xi)^T P) \\ & - \lambda_0(\Theta + \Phi)^2 \otimes PR^{-1}P] \Psi \\ & + \Psi^T [(M\Theta + M\Phi - a_1 M) \otimes PR^{-1}P] \Psi \\ \leq & \Psi^T [M(\Theta + \Phi) \otimes (PA^\xi + (A^\xi)^T P - PR^{-1}P)] \Psi \\ < & 0. \end{aligned} \quad (48)$$

Meanwhile, we select a_2 to satisfy the following inequality:

$$a_2 > \frac{\|P_i\|^2 \|Q_i^{-1}\| \|J_i^\delta\|^2}{\epsilon_{i1} m_i \lambda_{\min}(L_1^T L_1) \|P\|^2 \|R^{-1}\|}$$

then

$$\begin{aligned} \tilde{\xi}^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|J_i^\delta\|^2 - a_2 M L_1^T L_1 \otimes PR^{-1}P \right] \tilde{\xi} \\ < 0. \end{aligned} \quad (49)$$

It turns out from (46)–(49) that $\dot{V} < 0$, hence the error

$$\lim_{t \rightarrow \infty} \tilde{\xi}_i = 0, \lim_{t \rightarrow \infty} \psi_i = 0, \lim_{t \rightarrow \infty} \tilde{\delta}_i = 0, \lim_{t \rightarrow \infty} e_i = 0.$$

Therefore, Problem 1 is solved for systems (2)–(5) under the static graph.

(b) Now, we discuss the situation when $\varpi_i = 0$ and $G_s = G^{\sigma(t)}$. Under Assumption 2, graph $G^{\sigma(t)}$ and matrix $L_1^{\sigma(t)}$ are fixed on each interval $[t_k, t_{k+1})$. We chose the same V in (35) as pairwise Lyapunov function and the same (a_1, a_2) , where $\lambda_0 = \lambda_0^{\sigma(t)}$ and $L_1 = L_1^{\sigma(t)}$. Similar to the proof of part (a), we can obtain that $\dot{V} < 0$ on each interval $[t_k, t_{k+1})$. Therefore, each linear time-invariant subsystem of the following system is asymptotically stable:

$$\dot{\tilde{\xi}} = (\mathbf{I}_n \otimes A^\xi - (\Theta + \Phi)L_1^{\sigma(t)} \otimes R^{-1}P)\tilde{\xi} = A_c^{\sigma(t)}\tilde{\xi}. \quad (50)$$

It turns out that state matrices $A_c^{\sigma(t)}$ of each closed-loop subsystem are Hurwitz, and there exists a Lyapunov function $V_4 = \tilde{\xi}^T P^{\sigma(t)} \tilde{\xi}$ with quadratic form that satisfies

$$A_c^{\sigma(t)} P^{\sigma(t)} + P^{\sigma(t)} (A_c^{\sigma(t)})^T < 0 \quad (51)$$

on each interval $[t_k, t_{k+1})$. If we define the piecewise Lyapunov function for observer error as V_4 , and $A_c^{\sigma(t)} P^{\sigma(t)} + P^{\sigma(t)} (A_c^{\sigma(t)})^T = -W^{\sigma(t)}$, we can conclude that $\dot{V}_4 < -\lambda_{\min}(W^{\sigma(t)})V_4$ in each time interval, where $\lambda_{\min}(W^{\sigma(t)})$ represents the minimum eigenvalue of $W^{\sigma(t)}$. As $P^{\sigma(t)}$ is a switching matrix, there exists a constant $\mu \geq 1$, such that $P^k \leq \mu P^l$, $k, l \in \{1, 2, \dots, M\}$. Similar to the dwell-time-based works in [31], it follows that $\lim_{t \rightarrow \infty} \tilde{\xi}_i = 0$ if the average dwell time of each agent is longer than a positive threshold

$$\tau > \tau^* = \frac{\ln \mu}{\lambda_{\min}(W^{\sigma(t)})}. \quad (52)$$

Under Assumption 2, $L_1^{\sigma(t)}$ is a bounded symmetric positive definite matrix when $t \in [0, +\infty)$. In view of (13), Φ is not less than zero, and Θ is an increasing function greater than zero when $t \in [0, +\infty)$. We conclude that the threshold can be decreased by increasing the initial value of adaptive parameter Θ . Therefore, τ^* becomes sufficiently small by choosing a sufficiently large initial value of θ_i , that is, $\theta_i(0)$ for $i = 1, 2, \dots, n$.

It can also be verified by Barbalat's Lemma [33]. The piecewise Lyapunov function V_4 satisfies

$$V_4 \geq 0, \dot{V}_4 < 0$$

on each time interval. Similar to the analysis in [24], it follows from (16) that $\|\dot{\tilde{\xi}}\|$ is bounded over $[0, +\infty)$ by choosing a sufficiently large $\theta_i(0)$, and \ddot{V}_4 is also bounded. By Barbalat's Lemma

$$\lim_{t \rightarrow \infty} \tilde{\xi}_i = 0. \quad (53)$$

Note that V_1 can be designed as a common Lyapunov function candidate for δ_i because no switching signal $\sigma(t)$ is involved in V_1 . It follows from (53) that

$$\lim_{t \rightarrow \infty} \tilde{\delta}_i = 0, \lim_{t \rightarrow \infty} e_i = 0.$$

Thus, Problem 1 is solved under switching topologies. ■

The second result of this article is as follows.

Theorem 2: Consider systems (2)–(5) satisfying Assumptions 1–4. (a) When $\varpi_i \in L_2[0, \infty]$ and $G_s = G$, Problem 2 is solvable by distributed observer (13) and state feedback controller (22), if conditions (14) and (26) are satisfied.

(b) When $\varpi_i \in L_2[0, \infty]$ and $G_s = G^{\sigma(t)}$, Problem 2 is solvable by observer (13) with a sufficiently large $\theta_i(0)$ under condition (14) and controller (22) under condition (26).

Proof: (a) When $\varpi_i \in L_2[0, \infty]$ and $G_s = G$, the closed-loop error system with controller (22) is altered to

$$\dot{\tilde{\delta}}_i = \tilde{A}_i^c \tilde{\delta}_i + J_i^{\delta} \tilde{\xi} + E_i \varpi_i, e_i = C_i^c \tilde{\delta}_i, \quad i = 1, 2, \dots, n. \quad (54)$$

Design the Hamiltonian function with the same V in (35) as

$$\begin{aligned} J &= \dot{V} + \sum_{i=1}^n e_i^T e_i - \sum_{i=1}^n \gamma_i^2 \varpi_i^T \varpi_i \\ &\leq \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^{\delta})^T P_i + P_i \tilde{A}_i^{\delta} - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^{\delta} R_i^{-1} (\tilde{B}_i^{\delta})^T P_i \right. \\ &\quad \left. - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^{\delta} R_i^{-1} (\tilde{B}_i^{\delta})^T P_i + \epsilon_{i1} Q_i \right] \tilde{\delta}_i \end{aligned}$$

$$\begin{aligned} &+ \sum_{i=1}^n \varpi_i^T E_i^T P_i \tilde{\delta}_i + \sum_{i=1}^n \tilde{\delta}_i^T P_i E_i \varpi_i + \sum_{i=1}^n \tilde{\delta}_i^T C_i^c C_i^c \tilde{\delta}_i \\ &+ \Psi^T [M(\Theta + \Phi) \otimes (PA^{\xi} + (A^{\xi})^T P - PR^{-1}P)] \Psi \\ &+ \tilde{\xi}^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|J_i^{\delta}\|^2 \right. \\ &\quad \left. - a_2 M L_1^T L_1 \otimes PR^{-1}P \right] \tilde{\xi} - \sum_{i=1}^n \gamma_i^2 \varpi_i^T \varpi_i. \quad (55) \end{aligned}$$

Inequality (55) can be written as

$$\begin{aligned} J &\leq \Psi^T [M(\Theta + \Phi) \otimes (PA^{\xi} + (A^{\xi})^T P - PR^{-1}P)] \Psi \\ &+ \tilde{\xi}^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|J_i^{\delta}\|^2 \right. \\ &\quad \left. - a_2 M L_1^T L_1 \otimes PR^{-1}P \right] \tilde{\xi} \\ &+ \sum_{i=1}^n \tilde{\delta}_i^T \left[(\tilde{A}_i^{\delta})^T P_i + P_i \tilde{A}_i^{\delta} - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^{\delta} R_i^{-1} (\tilde{B}_i^{\delta})^T P_i \right. \\ &\quad \left. - \frac{1}{2\epsilon_{i1}} P_i \tilde{B}_i^{\delta} R_i^{-1} (\tilde{B}_i^{\delta})^T P_i + \frac{1}{\gamma_i^2} P_i E_i E_i^T P_i \right. \\ &\quad \left. + \frac{1}{\gamma_i^2} C_i^T C_i + \epsilon_{i1} Q_i \right] \tilde{\delta}_i \\ &- \sum_{i=1}^n \left(\varpi_i - \frac{1}{\gamma_i^2} E_i^T P_i \tilde{\delta}_i \right)^T \gamma_i^2 \left(\varpi_i - \frac{1}{\gamma_i^2} E_i^T P_i \tilde{\delta}_i \right). \quad (56) \end{aligned}$$

Integrating function J results in

$$\begin{aligned} \int_0^{\infty} J dt &= V(\infty) - V(0) + \sum_{i=1}^n \int_0^{\infty} (e_i^T e_i - \gamma_i^2 \varpi_i^T \varpi_i) dt \\ &\leq \int_0^{\infty} \Psi^T [M(\Theta + \Phi) \otimes (PA^{\xi} + (A^{\xi})^T P - PR^{-1}P)] \Psi dt \\ &+ \int_0^{\infty} \tilde{\xi}^T \left[\frac{1}{\epsilon_{i1}} \|P_i\|^2 \|Q_i^{-1}\| \|J_i^{\delta}\|^2 P \right. \\ &\quad \left. - a_2 M L_1^T L_1 \otimes PR^{-1}P \right] \tilde{\xi} dt \\ &- \sum_{i=1}^n \int_0^{\infty} \tilde{\delta}_i^T \epsilon_{i2} Q_i \tilde{\delta}_i dt \\ &- \sum_{i=1}^n \int_0^{\infty} \left(\varpi_i - \frac{1}{\gamma_i^2} E_i^T P_i \tilde{\delta}_i \right)^T \gamma_i^2 \left(\varpi_i - \frac{1}{\gamma_i^2} E_i^T P_i \tilde{\delta}_i \right) dt \\ &\leq 0. \quad (57) \end{aligned}$$

Since the conditions in (14) and (26) are satisfied, $\dot{V} < 0$ with the same value of a_1 and a_2 . The integral of Hamiltonian function meets $\int_0^{\infty} J dt \leq 0$. In view of (35) and (57), we have the bounded L_2 robust condition (58) holds for $i = 1, 2, \dots, n$

$$\int_0^{\infty} e_i^T e_i dt \leq \gamma_i^2 \int_0^{\infty} \varpi_i^T \varpi_i dt. \quad (58)$$

Therefore, Problem 2 is solved under the static graph.

(b) Similar to the analysis of Theorem 1(b), we can show that the strategies can be extended to switching graph by choosing a

sufficiently large $\theta_i(0)$. Note that the observer can still decouple the system dynamics from the network by selecting a large $\theta_i(0)$ under the switching signal by Barbalat's Lemma. This completes the proof. ■

Remark 4: By treating disturbances as separable into homogeneous and heterogeneous disturbances, Theorem 2 provides a solution for robust formation control problems under both types of disturbances.

Remark 5: The proposed formation control strategy differs from [7] and [22], in that it can automatically adapt to static graph without requiring information of communication Laplacian matrices. The strategy can be extended to the switching graph by choosing a sufficiently large $\theta_i(0)$ under Assumption 2. It is a model-based strategy with known A_0 and A_w . It should be mentioned that an observer has been proposed to estimate the state matrix of the exosystem in [21], but the knowledge of minimum eigenvalues of the Laplacian matrix is required. It is, however, technically challenging to remove both topology and model information in controller design. A solution that we are working on is to combine an adaptive observer with an online iterative reinforcement learning algorithm to realize fully distributed and model-free strategies [34].

Remark 6: Under Assumption 3 and the definition of p -copy internal model in (18), the pair (A_i^δ, B_i^δ) in (23) is stabilizable. Therefore, Riccati (23) admits a unique positive definite matrix P_i by given $Q_i = Q_i^T > 0$, $R_i = R_i^T > 0$, $\epsilon_{i1} > 0$, and $\epsilon_{i2} > 0$ [35]. Similarly, there exists a unique definite matrix P_i satisfying Riccati (26) by choosing parameters as $Q_i = Q_i^T > 0$, $R_i = R_i^T > 0$, $\epsilon_{i1} > 0$, $\epsilon_{i2} > 0$, and $0 < \gamma_i < 1$. For each agent, γ_i in (9) is a robust index, which indicates that the ratio of the L_2 function of output error to heterogeneous disturbances is less than or equal to γ_i . When the index is less than 1 and closer to zero, the system is more robust to heterogeneous disturbances.

IV. SIMULATION VERIFICATION

This section presents the design of a tracking and patrolling formation for bushfire monitoring. Six heterogeneous unmanned aerial vehicles (UAVs) are deployed to perform the tracking and patrolling tasks under the proposed formation control strategy that is subject to a homogeneous main wind disturbance. Additional heterogeneous disturbances are introduced to further verify the performance of the bounded L_2 robust formation output controller.

A. Design of Tracking and Patrolling Formation

A generally accepted spread model for high-speed movement and high-density free-burning bushfire is the ellipse model [36]. In the wind coordinate system, the bushfire range can be described as an expanding ellipse after time t from the ignition point. With the homogeneous fuel bed and uniform terrain, the length-to-width ratio of the ellipse is proportional to the wind speed.

Following the wind direction and fire center provides a way to track and patrol the edge of the spreading bushfires [30]. The fire center can be regarded as a virtual leader with system state $x_0 = \text{col}(x_c, y_c, \dot{x}_c, \dot{y}_c)$ and output $y_0 = \text{col}(x_c, y_c)$, where (x_c, y_c) and (\dot{x}_c, \dot{y}_c) denote the position and velocity of the fire

center, respectively. Therefore, the system matrices are

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, C_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (59)$$

The initial state is given as $x_0 = [0 \ 1 \ 2 \ 1]^T$.

In the wind coordinate system, a TVF to patrol the bushfire boundary can be designed as follows:

$$\begin{aligned} \dot{f}_i^x &= at \cos(\rho t + d(\theta_i^0)) \\ \dot{f}_i^y &= bt \sin(\rho t + d(\theta_i^0)) \\ \dot{f}_i^x &= a \cos(\rho t + d(\theta_i^0)) - at\rho \sin(\rho t + d(\theta_i^0)) \\ \dot{f}_i^y &= b \sin(\rho t + d(\theta_i^0)) + bt\rho \cos(\rho t + d(\theta_i^0)) \end{aligned} \quad (60)$$

where $f_i = [f_i^x, f_i^y, \dot{f}_i^x, \dot{f}_i^y]^T$ represents the state variable of i th TVF, and ρ denotes the patrolling rate. Parameters a and b are the spreading rates of the major axis and the minor axis, respectively. Notation $d(\theta_i^0)$ is the initial mapping patrol angle based on the formation mapping algorithms derived in our previous study [7], [37].

Taking the reference frame into consideration, the system matrices in the inertial coordinate system becomes

$$\begin{aligned} A_i^f &= (I_2 \otimes D)A_w^f(I_2 \otimes D)^T, C_i^f = C_0 \\ A_w^f &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \rho^2 & 0 & 0 & -\frac{2a\rho}{b} \\ 0 & \rho^2 & \frac{2b\rho}{a} & 0 \end{bmatrix}, D = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \end{aligned} \quad (61)$$

where D is the direction cosine matrix (DCM) [38] from the wind coordinate system to the inertial coordinate system, and α is the rotation angle between the two coordinate systems. We choose the parameters as $a = 0.2$, $b = 0.1$, $c = [2 \ 1]^T$, $\rho = 1$, and $\alpha = 26.5^\circ$.

Remark 7: The system matrices of n reference formation models in (5) can differ or be identical for n agents. For example, the patrol angular rate ρ in (61) can be different for agents operating at different altitude planes. We assume that heterogeneous UAVs are flying at the same altitude plan with a common patrol rate in the given task. Therefore, A_i^f is the same for each agent. Based on the distributed observer design in [21], an agent designs its own formation reference model by observing the spread model information $\beta_i = \mu_i \sum_{j=0}^n a_{ij}(t)(\beta_j - \beta_i)$, where $\mu_i > 0$ and $\beta_i = \text{col}(\alpha_1, \alpha_2, \dots, \alpha_m)$ is the vector of minimal polynomial coefficients of A_i^f . From the feature of the internal model design in Lemma 3, the pair (A_i^δ, B_i^δ) with observed $\hat{\Sigma}_i^1$ in (22) is still stabilizable. Therefore, Riccati equations (23) and (26) always have solutions that choose parameters satisfying the same conditions as those in Remark 6 [21]. Furthermore, the position output deviation of the reference formation is achieved through the initial mapping patrol angle $d(\theta_i^0)$ from the formation matching algorithms based on deviation reference $\delta\theta_0 = \pi/3$. For the patrolling task, it is not necessary for the agents to be evenly distributed on the edge and there is no requirement to know the agent's total number.

Based on our previous study on scalable formation matching algorithms [7], [37], we know that each agent can determine its own optimal conflict-free mapping $d(\theta_i^0)$ based only on local information. Therefore, the design of the TVF model and the mapping of the initial state are realized in a distributed manner for each agent.

B. Robust Formation Control Under Uncertainties and Homogeneous Disturbance

A heterogeneous MAS is considered with six UAVs consisting of three Qball-X4 [39] and three Qball2 [7]. Based on the quadrotor models in [7], the latitudinal model of UAVs is simplified to the following linear fourth-order systems:

$$\begin{aligned} \begin{bmatrix} \dot{p}_i^x \\ \dot{v}_i^x \\ \dot{\theta}_i^x \\ \dot{q}_i \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^x \\ v_i^x \\ \theta_i^x \\ q_i \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_i^m L_i}{I_i^x} \end{bmatrix} u_i^x \\ &= A_i^x x_i^x + B_i^x u_i^x \end{aligned} \quad (62)$$

$$\begin{aligned} \begin{bmatrix} \dot{p}_i^y \\ \dot{v}_i^y \\ \dot{\phi}_i^y \\ \dot{p}_i \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^y \\ v_i^y \\ \phi_i^y \\ p_i \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_i^m L_i}{I_i^y} \end{bmatrix} u_i^y \\ &= A_i^y x_i^y + B_i^y u_i^y \end{aligned} \quad (63)$$

where g is the gravity acceleration. For the i th agent, p_i^x and p_i^y represent the horizontal positions in the x and y directions. Variables θ_i^x , ϕ_i^y , q_i , and p_i donate the pitch angle, roll angle, pitch angular rate, and roll angular rate, respectively. Parameters k_i^m and L_i are the force coefficient and the lever arm, and I_i^x and I_i^y indicate moments of inertia in x and y directions.

Expressing in terms of heterogeneous MAS (2), the system states and matrices of heterogeneous multi-UAV systems are

$$x_i = \text{col}(x_i^x, x_i^y) = \text{col}(p_i^x, v_i^x, \theta_i^x, q_i, p_i^y, v_i^y, \phi_i^y, p_i)$$

$$\tilde{A}_i = \text{blkdiag}(\tilde{A}_i^x, \tilde{A}_i^y) = \text{blkdiag}(A_i^x + \Delta A_i^x, A_i^y + \Delta A_i^y)$$

$$\tilde{B}_i = \text{blkdiag}(\tilde{B}_i^x, \tilde{B}_i^y) = \text{blkdiag}(B_i^x + \Delta B_i^x, B_i^y + \Delta B_i^y).$$

For $i = 1, 2, 3$, $k_i^m = 12\text{N}$, $L_i = 0.2\text{m}$, and $I_i^x = I_i^y = 0.03\text{kg}\cdot\text{m}^2$, while for $i = 4, 5, 6$, $k_i^m = 120\text{N}$, $L_i = 0.2\text{m}$, and $I_i^x = I_i^y = 0.03\text{kg}\cdot\text{m}^2$. The initial positions of the six agents are taken as $(0.5, 1)\text{m}$, $(-3, 2)\text{m}$, $(-4, 0)\text{m}$, $(-6, -1)\text{m}$, $(0, -3)\text{m}$, and $(-2, 0)\text{m}$, respectively. We take the initial velocities of these agents as $(-0.3, -0.1)\text{m/s}$, $(0, 0)\text{m/s}$, $(-3, -2)\text{m/s}$, $(-1, 2)\text{m/s}$, $(-0.3, 1)\text{m/s}$, and $(-1, 1)\text{m/s}$, respectively. The initial attitude angles of UAVs are $(0, 0)\text{rad}$, $(0, 0)\text{rad}$, $(0, -0.1)\text{rad}$, $(0, 0)\text{rad}$, $(0, 0)\text{rad}$, $(0, -0.1)\text{rad}$, and the angular rates are $(0.1, 0)\text{rad/s}$, $(0.2, -0.1)\text{rad/s}$, $(-0.2, -0.1)\text{rad/s}$, $(0, 0)\text{rad/s}$, $(0, 0)\text{rad/s}$, and $(0, 0)\text{rad/s}$, respectively. Note that the dynamic in z direction of UAVs is not considered here. They are assumed to fly on the same height plane with an altitude of 10 m.

Taking into account load changes and the unmodeled dynamics, the following system uncertainties are included in this

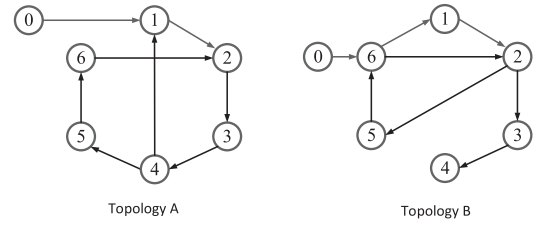


Fig. 1. Switching topologies of the heterogeneous MAS.

multi-UAV system:

$$\begin{aligned} \Delta A_i^k &= \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \Delta B_i^k = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, i = 1, 2, 3, k = x, y \\ \Delta A_i^k &= \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \Delta B_i^k = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, i = 4, 5, 6, k = x, y. \end{aligned}$$

We treat the main wind disturbance as homogeneous disturbance (3) with $A_\omega = 0$, $\omega_0 = [\omega_0^x, \omega_0^y]^T = [2, 1]$. The homogeneous disturbance input matrices and uncertainties are chosen as

$$D_i^k = [1 \ 1 \ 1 \ 1]^T, \Delta D_i^k = [0 \ 0 \ 0.5 \ 0]^T$$

$$\tilde{D}_i = \text{blkdiag}(\tilde{D}_i^x, \tilde{D}_i^y) = \text{blkdiag}(D_i^x + \Delta D_i^x, D_i^y + \Delta D_i^y).$$

Shown in Fig. 1 is the communication network of the MAS that switches between topology A and topology B per second. Note that both topologies A and B have a spanning tree with node 0 as the root. Matrix R in the distributed adaptive observer in (13) and (14) is set to $0.5I$, and the initial variables $\theta_i(0) = 10$. Taking into consideration the homogeneous disturbance and system uncertainties, we choose the parameters of robust controller in (22)–(24) as $\epsilon_1 = [0.05, 0.1, 0.1, 0.1, 0.1, 0.1]$, $\epsilon_2 = [1, 1, 1, 2, 2, 2]$ and $R_i = Q_i = I$ for $i = 1, 2, \dots, n$.

The simulation results of the observation errors and formation control of MAS are shown in Figs. 2 and 3, respectively.

We observe from Fig. 2 that the distributed adaptive observer of each agent tracks the positions of fire center (x_c, y_c) and its velocities (v_x, v_y) within 10 s under switching topologies. The observation errors of the homogeneous wind disturbance converge to zero in 10 s. The 3 and 2D positions of agents are updated every 10 s in Fig. 3. The trajectory of agent 1 is also provided in Fig. 3. It can clearly be seen that the agents can track the bushfire center and patrol along the edge of the spreading ellipse. Regarding the formation errors of MAS (2) with uncertainties and homogeneous disturbance when using the adaptive observer in (13) and (14) and robust formation controllers presented in (22)–(24), the errors converge to zero within 25 s under switching topologies, as shown in Fig. 3.

C. Robust Formation Control Subject to Uncertainties and Homogeneous/Heterogeneous Disturbances

When additional heterogeneous disturbances are considered, the disturbance input matrix is $E_i^k = [0 \ 0 \ 0 \ 20]$ for

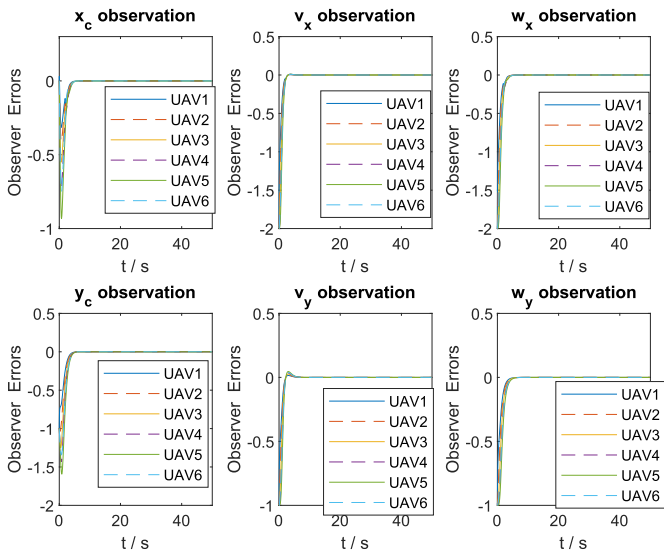


Fig. 2. Observation errors for the fire center and wind disturbance.

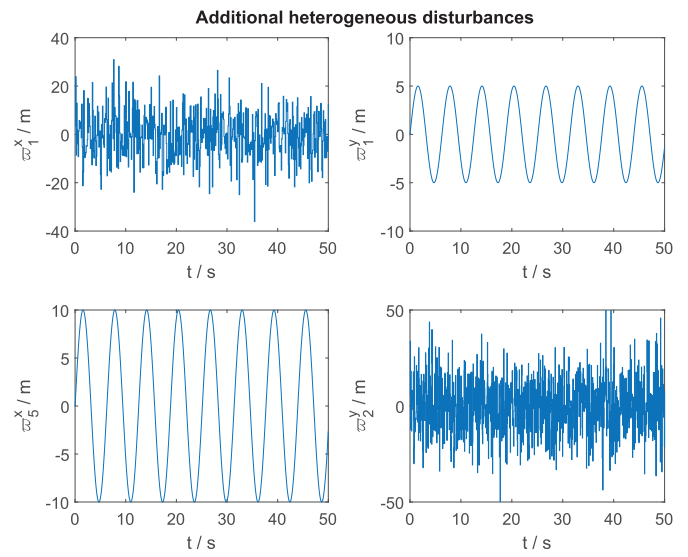


Fig. 4. Additional heterogeneous disturbances of low sinusoidal frequency and band-limited noise for performance verification.

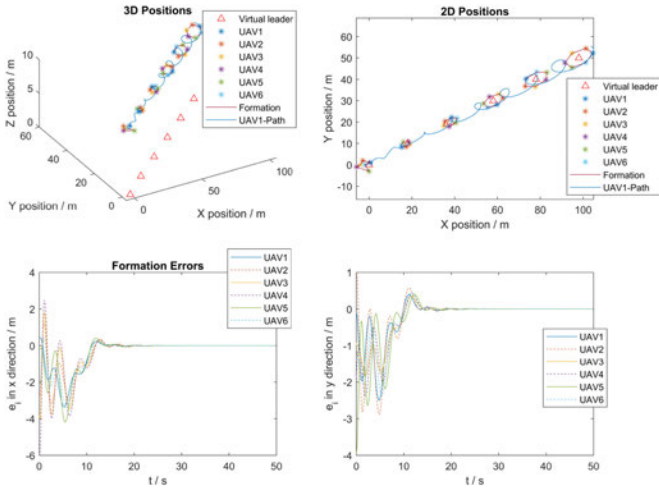


Fig. 3. TVF of the heterogeneous MAS subject to system uncertainties and homogeneous disturbance.

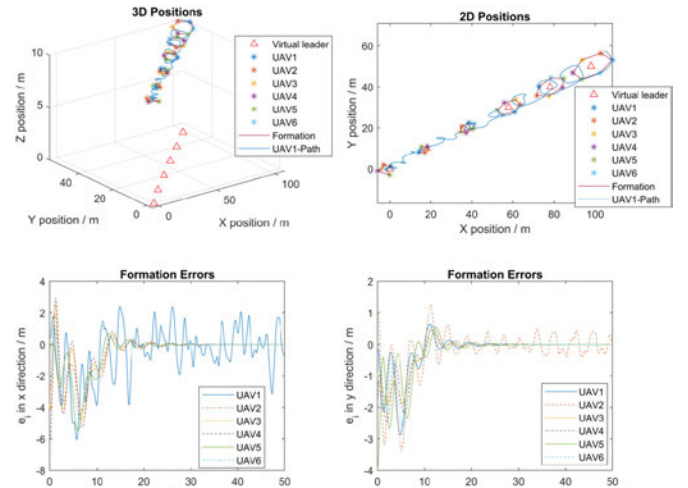


Fig. 5. TVF of the heterogeneous MAS under system uncertainties and homogeneous/heterogeneous disturbances by the method in [23].

$i = 1, 2, \dots, 6, k = x, y$. As shown in Fig. 4, a band-limited white noise w_1^x with 10-dB/Hz noise power and 0.1 s sample time, and a bounded noise $w_1^y = 5 \sin(t)$ are added to the x channel and the y channel of the first agent, respectively. Another band-limited white noise w_2^y with 10 dB/Hz noise power and 0.05 s sample time is added to the y channel of the second agent. Additional noise added to the x channel of the fifth agent is $w_5^x = 10 \sin(t)$. The system parameters, initial states, uncertainty matrices, homogeneous disturbance and parameters of the distributed adaptive observer are chosen to be the same as above. To illustrate the robustness of the proposed L_2 controller for multiple disturbances, comparative simulations are conducted with the robust output regulation method in [23]. The results using the robust output regulation method in [23] are shown in Fig. 5. Comparatively, the robust index in (26) of L_2 robust controller in (22) and (24) is set as $\gamma = [0.05, 0.1, 0.1, 0.1, 0.1, 0.1]$. The simulation results under the L_2 controller are shown in Fig. 6.

Fig. 5 shows that there are large fluctuations of errors in the x direction of the first agent and in the y direction of the second

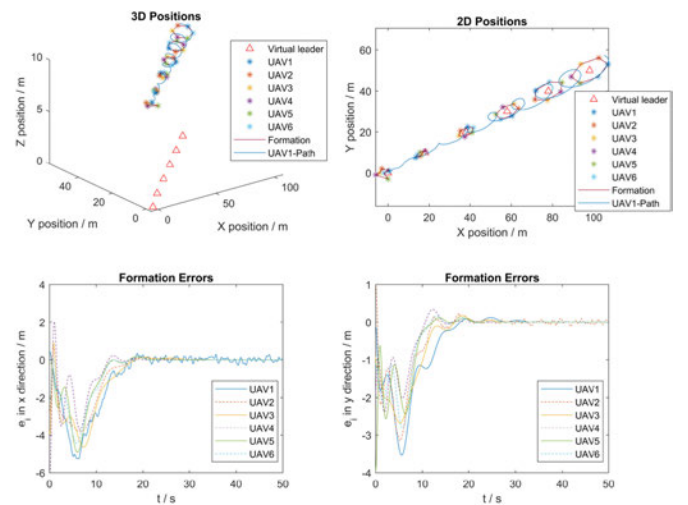


Fig. 6. TVF of the heterogeneous MAS subject to system uncertainties and homogeneous/heterogeneous disturbances under the proposed controller.

agent by the method in [23], under additional heterogeneous disturbances. After applying the proposed L_2 robust formation controller in system (2) with the same additional heterogeneous disturbances, the formation errors almost converge to zero after 25 s in Fig. 6. The amplitudes of the error fluctuation have been greatly reduced compared with that in Fig. 5. Note that a smaller γ_i means that the i th agent has stronger abilities to resist additional heterogeneous disturbances. For the bounded white noise interference ϖ_1^x and ϖ_2^y , the formation errors on the corresponding channels are also bounded. For disturbances $\varpi_1^y = 5 \sin(t)$ and $\varpi_5^x = 10 \sin(t)$, the formation errors on the corresponding channels converge to zero. This is because the internal model dynamics of n extended exosystems in (20) contain the internal model dynamics of the ϖ_1^y and ϖ_5^x , so the error-free formation control can be realized on these channels. Furthermore, the design decouples the heterogeneous dynamic from communication topologies through distributed adaptive observers and internal model-based formation controllers, it improves the robustness of the entire system to individual uncertainties or failures. Figs. 5 and 6 show that even if agents 1, 2, and 5 are disturbed, other agents will not be affected and can continue to work under switching topologies.

V. CONCLUSION

In this article, a distributed robust formation control strategy is proposed for an uncertain heterogeneous multiagent system with multiple disturbances. A distributed observer is developed to estimate the virtual leader and homogeneous disturbance under static and switching topologies. Comparative simulation studies on the bushfire edge tracking and patrolling verified the effectiveness and robustness of the proposed strategy. Our future work will consider the nonlinear formation control problems for unified heterogeneous MASs based on fuzzy adaptive control [19], [40] and model-free deep reinforcement learning.

REFERENCES

- [1] W. Zou, Y. Huang, C. K. Ahn, and Z. Xiang, "Containment control of linear multiagent systems with stochastic disturbances via event-triggered strategies," *IEEE Syst. J.*, vol. 14, no. 4, pp. 4810–4819, Dec. 2020.
- [2] P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 1, pp. 161–175, Jan. 2021.
- [3] Y. Liu, P. Shi, H. Yu, and C. C. Lim, "Event-triggered probability-driven adaptive formation control for multiple elliptical agents," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published, doi: [10.1109/TSMC.2020.3026029](https://doi.org/10.1109/TSMC.2020.3026029).
- [4] X. Dong and G. Hu, "Time-varying formation control for general linear multi-agent systems with switching directed topologies," *Automatica*, vol. 73, pp. 47–55, 2016.
- [5] X. Dong, B. Yu, Z. Shi, and Y. Zhong, "Time-varying formation control for unmanned aerial vehicles: Theories and applications," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 1, pp. 340–348, Jan. 2015.
- [6] X. Dong, Y. Zhou, Z. Ren, and Y. Zhong, "Time-varying formation tracking for second-order multi-agent systems subjected to switching topologies with application to quadrotor formation flying," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5014–5024, Jun. 2017.
- [7] B. Yan, P. Shi, C.-C. Lim, C. Wu, and Z. Shi, "Optimally distributed formation control with obstacle avoidance for mixed-order multi-agent systems under switching topologies," *IET Control Theory Appl.*, vol. 12, no. 13, pp. 1853–1863, 2018.
- [8] Y. Guan, Z. Ji, L. Zhang, and L. Wang, "Controllability of heterogeneous multi-agent systems under directed and weighted topology," *Int. J. Control*, vol. 89, no. 5, pp. 1009–1024, 2016.
- [9] B. Yan, C. Wu, and P. Shi, "Formation consensus for discrete-time heterogeneous multi-agent systems with link failures and actuator/sensor faults," *J. Franklin Inst.*, vol. 356, no. 12, pp. 6547–6570, 2019.
- [10] F. Xiao and T. Chen, "Adaptive consensus in leader-following networks of heterogeneous linear systems," *IEEE Control Netw. Syst.*, vol. 5, no. 3, pp. 1169–1176, Sep. 2018.
- [11] Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and Y. Shen, "Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control," *Automatica*, vol. 93, pp. 26–32, 2018.
- [12] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. vol. 8 Philadelphia, PA, USA: SIAM, 2004.
- [13] H. Zhang, Y. Cai, Y. Wang, and H. Su, "Adaptive bipartite event-triggered output consensus of heterogeneous linear multiagent systems under fixed and switching topologies," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4816–4830, Nov. 2020.
- [14] S. Zuo and D. Yue, "Resilient output formation containment of heterogeneous multigroup systems against unbounded attacks," *IEEE Trans. Cybern.*, to be published, doi: [10.1109/TCYB.2020.2998333](https://doi.org/10.1109/TCYB.2020.2998333).
- [15] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader-follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, 2008.
- [16] M. Saim, S. Ghapani, W. Ren, K. Munawar, and U. M. Al-Saggaf, "Distributed average tracking in multi-agent coordination: Extensions and experiments," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2428–2436, Sep. 2018.
- [17] Y. Liu, H. Yu, P. Shi, and C.-C. Lim, "Formation control and collision avoidance for a class of multi-agent systems," *J. Franklin Inst.*, vol. 356, no. 10, pp. 5395–5420, 2019.
- [18] W. Jiang, G. Wen, Z. Peng, T. Huang, and R. A. Rahmani, "Fully distributed formation-containment control of heterogeneous linear multi-agent systems," *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3889–3896, Sep. 2019.
- [19] Y. Li, F. Qu, and S. Tong, "Observer-based fuzzy adaptive finite-time containment control of nonlinear multiagent systems with input delay," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 126–137, Jan. 2021.
- [20] H. Cai, F. L. Lewis, G. Hu, and J. Huang, "The adaptive distributed observer approach to the cooperative output regulation of linear multi-agent systems," *Automatica*, vol. 75, pp. 299–305, 2017.
- [21] H. Cai and J. Huang, "Output based adaptive distributed output observer for leader-follower multiagent systems," *Automatica*, vol. 125, 2021, Art. no. 109413.
- [22] T. Han and W. X. Zheng, "Bipartite output consensus for heterogeneous multi-agent systems via output regulation approach," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 68, no. 1, pp. 281–285, Jan. 2021.
- [23] Z. Li, M. Z. Chen, and Z. Ding, "Distributed adaptive controllers for cooperative output regulation of heterogeneous agents over directed graphs," *Automatica*, vol. 68, pp. 179–183, 2016.
- [24] B. Cheng, X. Wang, and Z. Li, "Event-triggered consensus of homogeneous and heterogeneous multiagent systems with jointly connected switching topologies," *IEEE Trans. Cybern.*, vol. 49, no. 12, pp. 4421–4430, Dec. 2019.
- [25] S. Li, J. Zhang, X. Li, F. Wang, X. Luo, and X. Guan, "Formation control of heterogeneous discrete-time nonlinear multi-agent systems with uncertainties," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4730–4740, Jun. 2017.
- [26] R. Yang, H. Zhang, G. Feng, and H. Yan, "Distributed event-triggered adaptive control for cooperative output regulation of heterogeneous multi-agent systems under switching topology," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4347–4358, Sep. 2018.
- [27] T. Başar and P. Bernhard, *H ∞ Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Berlin, Germany: 2008.
- [28] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.
- [29] G. Wen, W. Yu, Z. Li, X. Yu, and J. Cao, "Neuro-adaptive consensus tracking of multiagent systems with a high-dimensional leader," *IEEE Trans. Cybern.*, vol. 47, no. 7, pp. 1730–1742, Jul. 2017.
- [30] B. Yan, P. Shi, and C.-C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," *IEEE Trans. Automat. Sci. Eng.*, to be published, doi: [10.1109/TASE.2021.3103877](https://doi.org/10.1109/TASE.2021.3103877).
- [31] Y. Hua, X. Dong, J. Wang, Q. Li, and Z. Ren, "Time-varying output formation tracking of heterogeneous linear multi-agent systems with multiple leaders and switching topologies," *J. Franklin Inst.*, vol. 356, no. 1, pp. 539–560, 2019.
- [32] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton, NJ: Princeton Univ. Press, 2009.
- [33] B. Farkas and S. A. Wegner, "Variations on Barbălat's lemma," *Amer. Math. Monthly*, vol. 123, no. 8, pp. 825–830, 2016.

- [34] B. Yan, P. Shi, C.-C. Lim, and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," *Int. J. Robust Nonlinear Control*, to be published, doi: [10.1002/rnc.5828](https://doi.org/10.1002/rnc.5828).
- [35] V. Kucera, "A contribution to matrix quadratic equations," *IEEE Trans. Autom. Control*, vol. 17, no. 3, pp. 344–347, Jun. 1972.
- [36] G. Pery, "Current approaches to modelling the spread of wildland fire: A review," *Prog. Phys. Geogr.*, vol. 22, no. 2, pp. 222–245, 1998.
- [37] H. Yu, P. Shi, and C. C. Lim, "Scalable formation control in stealth with limited sensing range," *Int. J. Robust Nonlinear Control*, vol. 27, no. 3, pp. 410–433, 2017.
- [38] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of Unmanned Aerial Vehicles*. Dordrecht, The Netherlands: Springer, 2015.
- [39] F. Pan, L. Liu, and D. Xue, "Optimal PID controller design with kalman filter for Qball-X4 quad-rotor unmanned aerial vehicle," *Trans. Inst. Meas. Control*, vol. 39, no. 12, pp. 1785–1797, 2017.
- [40] S. Tong, X. Min, and Y. Li, "Observer-based adaptive fuzzy tracking control for strict-feedback nonlinear systems with unknown control gain functions," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3903–3913, Sep. 2020.



Cheng-Chew Lim (Life Senior Member, IEEE) received the B.Sc. degree (Hons.) in electronic and electrical engineering, and Ph.D. degree in electronic and electrical engineering from Loughborough University, Leicestershire, U.K., in 1977 and 1981, respectively.

He is currently a Professor with the University of Adelaide, Adelaide, Australia. His research interests include control and systems theory, autonomous systems, machine learning, and optimization techniques and applications.

Dr. Lim has served as an editorial board member for the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS and the *Journal of Industrial and Management Optimization*.



Bing Yan received the B.Sc. degree in automation and the M.Sc. degree in control theory and control engineering from the School of Automation, Northwestern Polytechnical University, Xi'an, China, in 2012 and 2015, respectively. She is currently working toward the Ph.D. degree in electrical and electronic engineering with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, Australia.

Her research interests include flight control, formation control, and multi-agent systems.



Chengfu Wu received the B.Sc. degree in automatic control of aircraft and the M.Sc. degree in navigation guidance and control from the School of Automation in Northwestern Polytechnical University, Xi'an, China, in 1983 and 1986 respectively.

He is currently a Professor with Northwestern Polytechnical University. His research interests include navigation guidance and flight control.



Peng Shi (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Newcastle, Callaghan, Australia, in 1994, the doctor of science degree from the University of Glamorgan, Wales, U.K., in 2006, and the doctor of engineering degree from the University of Adelaide, Australia, Adelaide, Australia, in 2015.

He is currently a Professor with the University of Adelaide. His research interests include systems and control theory with applications to network systems, cyber-physical systems, intelligent systems, and autonomous and robotic systems.

Dr. Shi is a Fellow of the Institution of Engineering and Technology and the Institution of Engineers Australia. He has served on the editorial board of a number of journals, including *Automatica*, IEEE TRANSACTIONS ON (AUTOMATIC CONTROL, *Cybernetics*, *Circuits and Systems*, and *Fuzzy Systems*), and IEEE CONTROL SYSTEMS LETTERS. He currently serves as the Vice President and Distinguished Lecturer of IEEE SMC Society.

Chapter 4

Robust Formation Control for Nonlinear Heterogeneous Multi-agent Systems Based on Adaptive Event-triggered Strategy

Statement of Authorship

Title of Paper	Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	B. Yan, P. Shi and C. -C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," IEEE Transactions on Automation Science and Engineering, doi: 10.1109/TASE.2021.3103877, 2021.

Principal Author

Name of Principal Author (Candidate)	Bing Yan		
Contribution to the Paper	Conceptualization, methodology, experiments, validation and writing-original draft		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25 Jul 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Peng Shi		
Contribution to the Paper	Polishing, checking and verification.		
Signature		Date	25 Jul 2022

Name of Co-Author	Cheng-Chew Lim		
Contribution to the Paper	Review, refine and validate		
Signature		Date	25/7/22

4.1 Introduction

Based on the observer-based decoupling idea from Chapter 3, a brand-new dual adaptive time-varying formations (TVF) control scheme is proposed in this chapter for nonlinear heterogeneous multi-agent systems (MAS) to deal with limited network bandwidth constraints. Compared with linear MAS, a more general system, unified nonlinear heterogeneous MAS, is considered subject to uncertainties and disturbances. To reduce the frequency of data transmission, a distributed dual adaptive event-triggered observer is presented for exosystem estimation, which removes the global communication information in both observer design and Zeno-free event-triggered strategy design while saving network resources. Then, a nonlinear p-copy internal model-based formation controller is designed with a dynamic distributed compensator for uncertainties and disturbances, which solves the robust heterogeneous TVF problem. Finally, both simulation and experiment are conducted for the tracking and patrolling formation of multiple vehicles. The results verify that the proposed scheme can significantly reduce communication frequency under the premise of ensuring the robustness of multi-agent formations.

4.2 Publication

B. Yan, P. Shi and C. -C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2021.3103877, 2021.

Robust Formation Control for Nonlinear Heterogeneous Multiagent Systems Based on Adaptive Event-Triggered Strategy

Bing Yan^{1b}, Peng Shi^{1b}, *Fellow, IEEE*, and Cheng-Chew Lim^{1b}, *Life Senior Member, IEEE*

Abstract—In this article, a distributed adaptive event-triggered formation control strategy is proposed for unified nonlinear heterogeneous multiagent systems under uncertainties and disturbances to achieve time-varying formations. To reduce the frequency of data transmission, a distributed dual adaptive observer with an event-triggered strategy is developed to estimate the states of a reference exosystem. Without incurring prior global information about a communication graph, a novel robust formation controller, with dynamic distributed compensators for uncertainties and disturbances, is designed based on an observer result and a nonlinear internal control principle. Finally, both simulation and experiment are conducted for tracking and patrolling formation to verify the effectiveness of the proposed formation control strategy and its robustness.

Note to Practitioners—This article addresses the collaborative formation problem of multiagent systems that has potential applications in transportation and disaster relief. The design of robust and energy-saving coordination strategies is challenging in heterogeneous multivehicle systems. The proposed distributed method is suitable for large-scale uncertain heterogeneous systems, and the use of the dual adaptive event-triggered strategy reduces the data transmission rate.

Index Terms—Adaptive observer, event-triggered strategy, heterogeneous nonlinear multiagent systems (MASs), robust formation control.

I. INTRODUCTION

IN UNCERTAIN environment, multiagent systems (MASs) consisting of agents with environmental awareness, communication, and self-organizing capabilities can not only overcome the limitations of single agent in terms of load, coverage, and fault tolerance but also improve the execution efficiency of cooperative tasks and survivability of whole systems [1], [2]. The consensus problem is widely studied for agents to reach an agreement, which is the fundamental problem of cooperative control of MASs [3]–[5]. As one of the typical applications

Manuscript received April 3, 2021; accepted August 4, 2021. This article was recommended for publication by Associate Editor X. Na and Editor H. Gao upon evaluation of the reviewers' comments. This work was supported in part by the Australian Research Council under Grant DP170102644. (Corresponding author: Peng Shi.)

The authors are with the School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia (e-mail: bing.yan@adelaide.edu.au; peng.shi@adelaide.edu.au; cheng.lim@adelaide.edu.au).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TASE.2021.3103877>.

Digital Object Identifier 10.1109/TASE.2021.3103877

of collaboration technology of MASs, formation control [6] is designed to arrange the spatial position of multiple agents to geometric patterns. The agents in a formation make their own decisions to reach consensus according to reference formations and adapt to cooperative tasks and uncertain environment. In the fields such as search reconnaissance, environmental monitoring, transportation, and disaster relief, formation control of MASs has potential application value due to the effective expansion of the scope of operation and reduction of the risks of human involvement under uncertain and hazardous environment.

The methods of formation control can be divided into the centralized methods [7], [8] and the distributed methods [9]–[12] based on control structures. For formation control with a centralized structure, there is a control center or a host to coordinate information transmission and arrange tasks. Although the structure is simple and easy to implement, it heavily depends on the control center and is vulnerable to a single point of failure. On the contrary, all the agents in a distributed structure are independent of each other, and decisions are made based only on local information of communication. Consequently, distributed formation control is more flexible and robust when the structure is centerless. However, most existing distributed control strategies are not fully distributed, because the design of control laws, especially for high-order MASs, relies on global communication information with prior knowledge. In practice, the global communication matrix information is time-varying and difficult to obtain for large-scale high-order MASs.

Compared with linear homogeneous agents, nonlinear heterogeneous agents are more flexible in task allocation according to different capabilities in cooperative operations, such as unmanned aerial vehicle–unmanned ground vehicle (UAV-UGV) collaboration MASs. One of the basic heterogeneous systems, hybrid-order MASs consisting of different order integrator systems, was first studied in [13]–[15]. However, the less restrictive heterogeneous systems with different orders and different dynamics are more common in practical applications. Inspired by the output regulation control theory for single system [16], the output regulation problems have attracted increasing attention for linear unified heterogeneous MASs [11], [17], [18] and nonlinear heterogeneous MASs [19], [20]. The work in [19] investigated a cooperative output regulation problem for a group of heterogeneous

networked systems with nonlinear uncertainties. A class of heterogeneous MASs comprised of several first-order nonlinear systems and second-order nonlinear systems was considered in [20]. However, issues on unified nonlinear heterogeneous MASs deserve more comprehensive investigations compared with some typical systems. In addition to output regulation problems, it remains a big challenge to resolve the conflict between formation consensus requirements and high-order nonlinear heterogeneous dynamics.

Furthermore, uncertain system parameters, external disturbances, and communication limitations affect the performance of control systems. Considerable efforts have been made for MASs under uncertainties and disturbances based on robust control [21]–[24]. The work in [21] provided a nonlinear robust control for homogeneous Euler–Lagrange systems to achieve formation containment against uncertainties. A formation control for discrete-time uncertain multivehicle systems was developed in [22] under disturbances. From a practical point of view, network constraints, such as limited communication bandwidth and data loss, should also be considered in the MAS design. Event-triggered control has been regarded as an effective way to decrease communication load [4], [5], [25] and robust to attacks/faults [18], [26]. In [4], a quantized event-triggered control was proposed for the consensus problem of homogeneous MASs with disturbances. The work in [25] studied the robust cooperative output regulation problem of linear heterogeneous MASs with additive disturbances via the celebrated internal model principle. However, theoretical challenges arise from the formation control of unified nonlinear heterogeneous MASs considering uncertainties, disturbances, and communication limitations at the same time. Furthermore, uncertain system parameters will also bring significant difficulties when designing event-triggered control strategies.

Motivated by the above observations, we have systematically studied the problem of robust formation control for heterogeneous MASs, and the main contributions in this article are as follows.

- 1) A time-varying formation control scheme is proposed for unified nonlinear heterogeneous MASs with different orders and dynamics under uncertainties and disturbances.
- 2) Compared with the works in [14], [18], and [19], the proposed strategy based on a brand-new dual adaptive observer and nonlinear internal model control principle is distributed without requiring any global information of communication topology.
- 3) A novel adaptive event-triggered strategy is designed by the Riccati equation approach to overcome network constraints and exclude the Zeno behavior. The strategy can be applied to actual platforms to significantly reduce the communication load.

The notation used in this article is standard. X^{-1} and Y^T represent the inverse of nonsingular and square matrix X and transpose of matrix Y , respectively. $\mathbf{0}$ is all-zero matrix with corresponding dimensions. Notation \otimes denotes the Kronecker product, and blkdiag represents the block diagonal

concatenation of matrix input arguments. For n vectors x_i , $i = 1, \dots, n$, $\text{col}(x_1, x_2, \dots, x_n) = [x_1^T, x_2^T, \dots, x_n^T]^T$.

II. PROBLEM FORMULATION

A heterogeneous nonlinear MASs with n agents can be modeled as

$$\begin{aligned} \dot{x}_i &= f_i(x_i, u_i, \omega, \Delta_i) \\ y_i &= h_i(x_i, u_i, \omega, \Delta_i), \quad i = 1, 2, \dots, n \end{aligned} \quad (1)$$

where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}$, and $y_i \in \mathbb{R}^{p_i}$ represent system states, input variables, and output variables of the i th agent, respectively. Functions f_i , h_i are sufficiently smooth nonlinear functions. The set of uncertain parameters is Δ_i . The dynamics of disturbance signal $\omega \in \mathbb{R}^{n_\omega}$ is described by

$$\dot{\omega} = A_\omega \omega. \quad (2)$$

Note that the linear system (2) we use to approximate disturbances with its dynamics is assumed known. Generally, the nonlinear function of disturbances can be expanded by the Fourier series and approximated as a combined signal via trigonometric functions. The state matrix A_ω can always be found from the trigonometric functions to predict changes of disturbance signals, especially for common disturbances, such as the main wind environment disturbances.

To implement the distributed formation control for heterogeneous MASs, a virtual leader is designed in the desired formation as

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 \\ y_0 &= C_0 x_0 \end{aligned} \quad (3)$$

where $x_0 \in \mathbb{R}^{n_0}$ and $y_0 \in \mathbb{R}^{p_0}$ are state variables and output variables of the virtual leader, respectively. The state matrix and output matrix of the system are A_0 and C_0 , respectively. Note that only some agents get the information from the virtual leader.

In addition to a virtual leader, the shape of formation is illustrated by time-varying formation (TVF) system

$$\begin{aligned} \dot{f}_i &= A_i^f f_i \\ y_i^f &= C_i^f f_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (4)$$

where $f_i \in \mathbb{R}^{n_f}$ and $y_i^f \in \mathbb{R}^{p_f}$ are the state and output of TVF, respectively. The matrices A_i^f and C_i^f are the state matrix and output matrix of TVF system, respectively.

Remark 1: As shown in Fig. 1, the output y_0 of system (3) represents the trajectory changes with time of the virtual leader, which determines the specific position of the whole formation when performing tasks. Outputs of system (4) represent the relative displacements of each agent relative to the virtual leader. For example, with different values of y_i^f from time t_1 to time t_3 , the reference formation changes from a diamond to a rectangle and then to a parallelogram in Fig. 1. Therefore, the dynamics of the virtual leader and TVF determines the desired movement process of MASs for adapting to different tasks.

We introduce the following definitions to develop our main results in Section III.

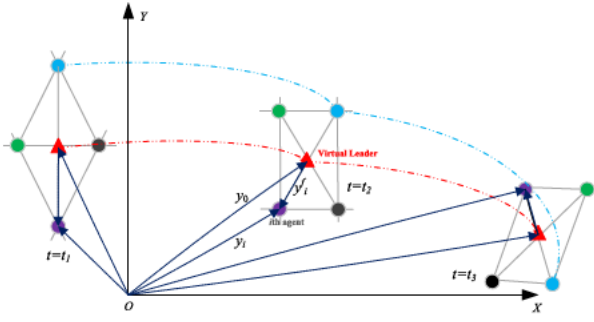


Fig. 1. TVF with a virtual leader.

Definition 1: We say that system (1) gradually reaches output consensus if

$$\lim_{t \rightarrow \infty} y_i - y_0 = 0, \quad i = 1, 2, \dots, n. \quad (5)$$

Definition 2: We say that system (1) with (3) and (4) gradually reaches formation consensus if

$$\lim_{t \rightarrow \infty} y_i - y_0 - y_i^f = 0, \quad i = 1, 2, \dots, n. \quad (6)$$

Remark 2: The difference between dynamics and state dimensions of MASs (1) makes it difficult to ensure that all states achieve the consensus. Assuming that the output dimensions of systems (1), (3), and (4) are the same, that is, $p_i = p_0 = p_f$, then the output consensus becomes possible. Inspired by the work in [16], the output consensus defined in Definition 1 is intended to lead outputs of heterogeneous systems to converge to the same value. The purpose of Definition 2 for MASs is to maintain the TVF shapes and track the specified trajectories.

To proceed with formation control design under system uncertainties and disturbances, we describe the robust control problems to be addressed as follows:

Problem 1: Nonlinear robust formation output control problem is to design a controller to meet condition (7) for system (1) with (3) and (4) when $\Delta_i \in W$, where W is open neighborhood of origin

$$\lim_{t \rightarrow \infty} e_i = 0, \quad i = 1, 2, \dots, n \quad (7)$$

where the formation error $e_i = h_i(x_i, u_i, \omega, \Delta_i) - y_0 - y_i^f$.

We end this section by introducing the following assumptions.

Assumption 1: Matrix $A_i^v = \text{blkdiag}(A_0, A_\omega, A_i^f)$ has no eigenvalues with negative real parts.

Assumption 2: Assume that the communication graph G of the heterogeneous MAS has a spanning tree, where the virtual leader is the globally reachable root node, and the subgraph G_1 except for the virtual leader is connected.

Assumption 3: Assume that $((\partial f_i)/(\partial x_i)(0, 0, 0, 0), (\partial f_i)/(\partial u_i)(0, 0, 0, 0))$ is stabilizable, and the pair $((\partial h_i)/(\partial x_i)(0, 0, 0, 0), (\partial f_i)/(\partial x_i)(0, 0, 0, 0))$ is detectable.

Assumption 4: For $i = 1, 2, \dots, n$ and $l = 1, 2, \dots,$

$$\text{rank} \begin{bmatrix} \frac{\partial f_i}{\partial x_i}(0, 0, 0, 0) - \lambda I & \frac{\partial f_i}{\partial u_i}(0, 0, 0, 0) \\ \frac{\partial h_i}{\partial x_i}(0, 0, 0, 0) & \frac{\partial h_i}{\partial u_i}(0, 0, 0, 0) \end{bmatrix} = n_i + p_i$$

for all λ given by

$$\lambda|\lambda = \sum_{l=1}^q l_i \lambda_l, \quad \sum_{l=1}^q l_l = l, \quad l_1, \dots, l_q = 0, 1, 2, \dots, l$$

where $\lambda_1, \dots, \lambda_q$ are the eigenvalues of A_i^v .

Assumption 5: There exist sufficiently smooth functions $x_i(v_i, \Delta_i)$ and $u_i(v_i, \Delta_i)$ with $x_i(0, 0) = 0$ and $u_i(0, 0) = 0$ satisfying

$$\begin{cases} \frac{\partial x_i(v_i, \Delta_i)}{\partial v_i} A_i^v v_i = f_i(x_i(v_i, \Delta_i), u_i(v_i, \Delta_i), v_i, \Delta_i) \\ 0 = h_i(x_i(v_i, \Delta_i), u_i(v_i, \Delta_i), v_i, \Delta_i) - y_0 - y_i^f \end{cases} \quad (8)$$

for $i = 1, 2, \dots, n$, $v_i \in V$ and $\Delta_i \in W$, where V and W are some open neighborhoods of the origins.

Remark 3: Note that Assumption 1 ensures that no state of the virtual leader, TVF, and disturbances will converge to zero automatically. Otherwise, the agents with the same convergent trajectory may collide, and the disturbances will automatically disappear. Therefore, Assumption 1 is reasonable to track a dynamic trajectory in a TVF without causing agent collisions under disturbances. Similar assumptions can be found in [19] and [27]. The existence of a common internal model system for nonlinear heterogeneous MASs to achieve formation output control is ensured by Assumptions 4 and 5, which are extended by the solvable conditions of output regulation equations for linear MASs [18] and nonlinear dynamic systems [16]. The virtual leader and heterogeneous agents considered for the formation control problem are all moving bodies that satisfy Newton's second theorem. Therefore, the existence of the common internal model dynamics by Assumptions 4 and 5 is reasonable in reality.

III. MAIN RESULTS

In this section, a dual distributed adaptive observer with an event-triggered mechanism is designed for an exosystem without prior knowledge of the global communication graph. Then, a robust formation control strategy is proposed for nonlinear heterogeneous MASs to solve Problem 1.

A. Dual Adaptive Observer With Event-Triggered Strategy

Since only parts of agents get information from the virtual leader, they share information through network communications. The distributed observer is necessary to estimate their states for applying to large-scale systems under robust distributed structures. An exosystem consisting of virtual leader and disturbance is recast as

$$\begin{cases} \dot{\xi} = A^\xi \xi \\ y^\xi = C^\xi \xi \end{cases} \quad (9)$$

where $\xi = \text{col}(x_0, \omega)$, $A^\xi = \text{blkdiag}(A_0, A_\omega)$, and $C^\xi = [C_0, 0]$.

Remark 4: The concept of exosystem is proposed in the theory of output regulation for a single plant [28]. It refers to an external generator of reference (and/or disturbance) signals for the output of the plant to track (and/or reject). The role of the exosystem in output regulation for a single

system has been thoroughly illustrated in [29]. The exosystem has been expanded to the output consensus problem of MASs to describe an external system that generating tracking signals and disturbance signals. For heterogeneous MASs, the exosystem represents an external common goal related to collaborative tasks. For example, the exosystem in (9) means that the common goal is to track the virtual leader under disturbances.

Error system of robust formation output control problem is rewritten as

$$e_i = y_i - y_i^c - y_i^f. \quad (10)$$

Based on Assumption 2, the virtual leader can be indexed by number 0, and the whole Laplacian matrix becomes

$$L_{n+1} = \begin{bmatrix} 0 & 0_{1 \times n} \\ L_0 & L_1 \end{bmatrix} \quad (11)$$

where $L_1 = L + A_{n0}$ and L is the Laplacian matrix of n heterogeneous MASs. The adjacency matrix between the virtual leader and the multiagent is $A_{n0} = \text{diag}[a_{i0}]$.

We design the distributed adaptive observer based on the event-triggered strategy as

$$\begin{aligned} \dot{\hat{\zeta}}_i &= A^\xi \hat{\zeta}_i - (R^{-1}P)\theta_i \psi_i \\ \psi_i &= \sum_{j=1}^n a_{ij} \left(e^{A^\xi(t-t_j^k)} \hat{\zeta}_i(t_j^k) - e^{A^\xi(t-t_i^k)} \hat{\zeta}_j(t_j^k) \right) \\ &\quad + a_{i0} e^{A^\xi(t-t_i^k)} (\hat{\zeta}_i(t_i^k) - \zeta(t_i^k)) \quad t \in [t_i^k, t_i^{k+1}) \\ \dot{\theta}_i &= \psi_i^T (PR^{-1}P) \psi_i \end{aligned} \quad (12)$$

where $\theta_i(0) > 1$. For a given positive definite matrices R and Q , symmetric positive definite matrix P is the solution to the equation

$$(A^\xi)^T P + PA^\xi - PR^{-1}P + Q = 0. \quad (13)$$

The sequence of time $\{t_i^0, t_i^1, \dots, t_i^k, \dots | i = 1, 2, \dots, n, k = 0, 1, \dots\}$ is the event-triggered instants of agent i . The time t_i^k is the latest time of agent j before current time t .

The measurement error of agent i can be constructed as

$$e_{\hat{\zeta}_i} = e^{A^\xi(t-t_i^k)} \hat{\zeta}_i(t_i^k) - \hat{\zeta}_i, \quad t \in [t_i^k, t_i^{k+1}). \quad (14)$$

We give the adaptive event-triggered strategy as

$$t_i^{k+1} = \inf\{t > t_i^k \mid \phi_i \theta_i e_{\hat{\zeta}_i}^T e_{\hat{\zeta}_i} \geq \psi_i^T \psi_i\} \quad (15)$$

where the adaptive parameter ϕ_i is updated by

$$\dot{\phi}_i = e_{\hat{\zeta}_i}^T \theta_i (PR^{-1}P) e_{\hat{\zeta}_i}, \quad \phi_i(0) > 1. \quad (16)$$

Defining the observer error variable as $\tilde{\zeta}_i(t) = \hat{\zeta}_i(t) - \zeta(t)$, where $\zeta(t) = e^{A^\xi(t-t_i^k)} \zeta(t_i^k)$, we denote related column vectors of n agents by

$$\begin{aligned} \underline{\zeta} &= I_n \otimes \zeta(t), \quad \tilde{\zeta} = \text{col}(\tilde{\zeta}_1(t), \tilde{\zeta}_2(t), \dots, \tilde{\zeta}_n(t)) \\ \hat{\zeta} &= \text{col}(\hat{\zeta}_1(t), \hat{\zeta}_2(t), \dots, \hat{\zeta}_n(t)) \\ \Psi &= \text{diag}(\psi_1, \psi_2, \dots, \psi_n), \quad \Theta = \text{diag}(\theta_1, \theta_1, \dots, \theta_n) \\ \Phi &= \text{diag}(\phi_1, \phi_2, \dots, \phi_n), \quad e_\xi = \text{col}(e_{\hat{\zeta}_1}, e_{\hat{\zeta}_2}, \dots, e_{\hat{\zeta}_n}). \end{aligned}$$

From the definition of Laplacian matrix L_{n+1} in (11), it follows that

$$\begin{aligned} \Psi &= (L_1 \otimes I) \hat{\zeta} + (L_0 \otimes I) \underline{\zeta} + (L_1 \otimes I) e_\xi \\ &= (L_1 \otimes I) \hat{\zeta} + (L_1 \otimes I) e_\xi. \end{aligned} \quad (17)$$

In view of (12), the derivatives of the variables are

$$\begin{aligned} \dot{\hat{\zeta}} &= (I_n \otimes A^\xi - \Theta L_1 \otimes R^{-1}P) \hat{\zeta} - (\Theta L_1 \otimes R^{-1}P) e_\xi \\ &= (I_n \otimes A^\xi) \hat{\zeta} - (\Theta \otimes R^{-1}P) \Psi \\ \dot{\Psi} &= (I_n \otimes A^\xi) \Psi \\ \dot{\Theta} &= \Psi^T (I_n \otimes PR^{-1}P) \Psi \\ \dot{\Phi} &= e_\xi^T (\Theta \otimes PR^{-1}P) e_\xi. \end{aligned} \quad (18)$$

Recall the following results in order to derive our main results.

Lemma 1 [30]: If the communication topology G contains a spanning tree and the virtual leader is the root node, the matrix L_1 in (11) is a nonsingular M-matrix. If the subgraph G_1 without the virtual leader is connected, the matrix L_1 is a nonsingular symmetric matrix.

Lemma 2 [16]: The pair of (Σ_1, Σ_2) is the p -copy internal model of square matrix A , if the pair has the form

$$\Sigma_1 = T \begin{bmatrix} S_1 & S_2 \\ 0 & \Sigma_1 \end{bmatrix} T^{-1}, \quad \Sigma_2 = T \begin{bmatrix} S_3 \\ \Sigma_2 \end{bmatrix} \quad (19)$$

where T is any nonsingular matrix, S_i , $i = 1, 2, 3$, are any constant matrices, and for any, $p > 0$,

$$\overline{\Sigma}_1 = \text{blkdiag}(a_{11}, \dots, a_{1p}), \quad \overline{\Sigma}_2 = \text{blkdiag}(a_{21}, \dots, a_{2p}).$$

If the minimal polynomial of A is

$$\lambda^n + a_1 \lambda^{(n-1)} + \dots + a_{(n-1)} \lambda + a_n$$

then (a_{1j}, a_{2j}) is controllable in the form of

$$\alpha_{1j} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -a_n & -a_{(n-1)} & \dots & -a_1 \end{bmatrix}, \quad \alpha_{2j} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

where $j = 1, 2, \dots, p$.

One of our main results of this article is given as follows.

Theorem 1: Consider systems (1), (3), and (4) satisfying Assumptions 1–5. The observer error of distributed dual adaptive observer (12) converges to zero under event-triggered mechanism (15) and (16), if condition (13) is met. Furthermore, the Zeno behavior can be excluded.

Proof: We define the Lyapunov functions as

$$V = V_1 + V_2 + V_3 \quad (20)$$

where

$$V_1 = \tilde{\zeta}^T (L_1^T \otimes P) \tilde{\zeta} \quad (21)$$

$$V_2 = \sum_{i=1}^n \frac{1}{2} (\theta_i - a_1)^2 \quad (22)$$

$$V_3 = \sum_{i=1}^n \frac{1}{2} (\phi_i - a_2)^2 \quad (23)$$

where a_1 and a_2 are constants to be designed later. According to the design of observer (12) and adaptive law (16), the following conditions hold:

$$\theta_i > 1, \quad \phi_i > 1, \quad V > 0.$$

Differentiating V_1 along the trajectory of (21) based on Lemma 1 gives

$$\begin{aligned} \dot{V}_1 \leq & \zeta^T \left[L_1^T \otimes \left(PA^\xi + (A^\xi)^T P \right) - 2L_1^T \Theta L_1 \otimes \Pi \right] \tilde{\zeta} \\ & - e_\xi^T (L_1^T \Theta L_1 \otimes \Pi) \tilde{\zeta} - \zeta^T (L_1^T \Theta L_1 \otimes \Pi) e_\xi \end{aligned} \quad (24)$$

where $\Pi = PR^{-1}P$. Substituting (18) into derivation of V_2 in (22) gives

$$\begin{aligned} \dot{V}_2 = & \sum_{i=1}^n (\theta_i - a_1) \dot{\theta}_i \\ = & \Psi^T ((\Theta - (a_1 - 1)I_n) \otimes \Pi) \Psi - \Psi^T (I_n \otimes \Pi) \Psi \\ \leq & \zeta^T (L_1^T \Theta L_1 \otimes \Pi) \tilde{\zeta} + e_\xi^T (L_1^T \Theta L_1 \otimes \Pi) \tilde{\zeta} \\ & + \zeta^T (L_1^T \Theta L_1 \otimes \Pi) e_\xi + e_\xi^T (L_1^T \Theta L_1 \otimes \Pi) e_\xi \\ & - \frac{(a_1 - 1)}{2} \zeta^T (L_1^T L_1 \otimes \Pi) \tilde{\zeta} \\ & + (a_1 - 1) e_\xi^T (L_1^T L_1 \otimes \Pi) e_\xi - \Psi^T (I_n \otimes \Pi) \Psi. \end{aligned} \quad (25)$$

Under the event-triggered schemes in (15) and (16), taking the derivation of V_3 along a trajectory in (23) implies

$$\begin{aligned} \dot{V}_3 = & \sum_{i=1}^n (\phi_i - a_2) \dot{\phi}_i = \sum_{i=1}^n e_{\xi_i}^T (\phi_i - a_2) \theta_i (PR^{-1}P) e_{\xi_i} \\ \leq & \Psi^T (I_n \otimes \Pi) \Psi - e_{\xi}^T (a_2 \Theta \otimes \Pi) e_{\xi}. \end{aligned} \quad (26)$$

Therefore, the derivation of Lyapunov function V is

$$\begin{aligned} \dot{V} = & \dot{V}_1 + \dot{V}_2 + \dot{V}_3 \\ \leq & \Psi^T \left[L_1^T \otimes \left(PA^\xi + (A^\xi)^T P \right) - L_1^T \Theta L_1 \otimes \Pi \right. \\ & \left. - \frac{(a_1 - 1)}{2} L_1^T L_1 \otimes \Pi \right] \Psi \\ & + e_\xi^T [(L_1^T \Theta L_1 - a_2 \Theta) \otimes \Pi + (a_1 - 1) L_1^T L_1 \otimes \Pi] e_\xi \\ \leq & \Psi^T \left[L_1^T \otimes \left(PA^\xi + (A^\xi)^T P \right) - L_1^T L_1 \otimes \Pi \right. \\ & \left. - \frac{(a_1 - 1)}{2} L_1^T L_1 \otimes \Pi \right] \Psi \\ & + e_\xi^T [(L_1^T \Theta L_1 - a_2 \Theta + (a_1 - 1) L_1^T \Theta L_1) \otimes \Pi] e_\xi \\ \leq & \Psi^T \left[L_1^T \otimes \left(PA^\xi + (A^\xi)^T P \right) \right. \\ & \left. + L_1^T \lambda_{\min}(L_1) \frac{(a_1 + 1)}{2} \otimes (-\Pi) \right] \Psi \\ & + e_\xi^T [(-a_2 + a_1 \lambda_{\max}(L_1^T L_1)) \Theta \otimes \Pi] e_\xi. \end{aligned} \quad (27)$$

Choosing a_1 and a_2 to satisfy

$$a_1 > \frac{2}{\lambda_{\min}(L_1)} - 1, \quad a_2 > a_1 \lambda_{\max}(L_1^T L_1) \quad (28)$$

then we have

$$\dot{V} \leq \Psi^T \left[L_1^T \otimes \left(PA^\xi + (A^\xi)^T P - \Pi \right) \right] \Psi. \quad (29)$$

From (13), it turns out that $\dot{V} < 0$, and hence, the observer error

$$\lim_{t \rightarrow \infty} \tilde{\xi}_i = 0.$$

Furthermore, the adaptive variables ψ_i and ϕ_i are bounded.

We now prove that the Zeno behavior can be excluded. Under event-triggered condition (15), taking the derivations of $\phi_i \theta_i e_{\xi_i}^T e_{\xi_i}$ and $\psi_i^T \psi_i$

$$\begin{aligned} \frac{d(\phi_i \theta_i e_{\xi_i}^T e_{\xi_i})}{dt} = & \dot{\phi}_i \theta_i e_{\xi_i}^T e_{\xi_i} + \phi_i \dot{\theta}_i e_{\xi_i}^T e_{\xi_i} \\ & + 2\phi_i \theta_i e_{\xi_i}^T (A^\xi e_{\xi_i} - \theta_i W_P \psi_i) \\ \leq & (\dot{\phi}_i + \dot{\theta}_i + 2A^\xi + 1) \phi_i \theta_i e_{\xi_i}^T e_{\xi_i} \\ & + (\phi_i \theta_i^3 \|W_P\|^2) \psi_i^T \psi_i \end{aligned} \quad (30)$$

$$\frac{d(\psi_i^T \psi_i)}{dt} = 2A^\xi \psi_i^T \psi_i \quad (31)$$

where $W_P = R^{-1}P$ and $\|W_P\|$ is the norm of matrix W_P . Defining function $J_i = (\phi_i \theta_i e_{\xi_i}^T e_{\xi_i}) / (\psi_i^T \psi_i)$ and taking the derivation of J_i

$$\begin{aligned} \dot{J}_i \leq & \frac{d(\phi_i \theta_i e_{\xi_i}^T e_{\xi_i})}{dt} \psi_i^T \psi_i - (\phi_i \theta_i e_{\xi_i}^T e_{\xi_i}) \frac{d(\psi_i^T \psi_i)}{dt} \\ & \frac{\psi_i^T \psi_i \psi_i^T \psi_i}{\psi_i^T \psi_i \psi_i^T \psi_i} \\ \leq & (\dot{\phi}_i + \dot{\theta}_i + 1) J_i + (\bar{\phi}_i \bar{\theta}_i^3 \|W_P\|^2) \end{aligned} \quad (32)$$

where the adaptive variables are bounded by $\phi_i \leq \bar{\phi}_i$ and $\theta_i \leq \bar{\theta}_i$. The solution of function J_i is

$$J_i \leq \frac{(\bar{\phi}_i \bar{\theta}_i^3 \|W_P\|^2)}{(\dot{\phi}_i + \dot{\theta}_i + 1)} \left(e^{(\dot{\phi}_i + \dot{\theta}_i + 1)\tau} - 1 \right) \quad (33)$$

where $\tau \leq t_i^{k+1} - t_i^k$ is the smallest time interval and meets the critical condition of (15)

$$\tau = \frac{1}{(\dot{\phi}_i + \dot{\theta}_i + 1)} \ln \left(\frac{(\dot{\phi}_i + \dot{\theta}_i + 1)}{(\bar{\phi}_i \bar{\theta}_i^3 \|W_P\|^2)} + 1 \right) > 0.$$

Therefore, the Zeno behavior can be excluded, which completes the proof. \square

Remark 5: Compared with the observers designed for a single uncertain system in [31] and [32], the local interaction information with an event-triggered mechanism of MASs is considered in observer (12) for estimating the states of the exosystem and overcoming network constraints. It should be emphasized that although the design of parameters a_1 and a_2 in (28) is based on the minimum eigenvalues of the global Laplace matrix L_1 , they only appear in the Lyapunov function instead of the design of the observer.

B. Nonlinear Robust Formation Output Control

In this section, the robust formation output control is designed for system (1) under uncertainties and disturbances based on the observer in (12).

Since the output of TVF (4) is different for each agent, the n exosystems considering the TVF dynamics are

$$\dot{v}_i = A_i^v v_i, \quad y_i^v = S_i v_i \quad (34)$$

where $v_i = \text{col}(\zeta, f_i)$, $A_i^v = \text{blkdiag}(A_i^\zeta, A_i^f)$, and $S_i = [C_i^\zeta, C_i^f]$. The augmented system of (1) and (34) with the formation output error in (7) is transformed to

$$\begin{aligned}\dot{x}_i &= f_i(x_i, u_i, v_i, \Delta_i) \\ \dot{v}_i &= A_i^v v_i \\ e_i &= h_i(x_i, u_i, v_i, \Delta_i) - S_i v_i.\end{aligned}\quad (35)$$

We give the following notations to approximate the nonlinear functions of system (35) by the Taylor expansion series approach in [16]. For any matrix A , $A^{(n)}$ is define as the Kronecker product of n A

$$A^{(0)} = 1, \quad A^{(1)} = A, \quad A^{(n)} = A \otimes A \cdots \otimes A. \quad (36)$$

For a vector $v = \text{col}(v_1, v_2, \dots, v_q)$, notation $v^{[l]}$ denotes

$$v^{[l]} = [v_1^l, v_1^{l-1}v_2, \dots, v_1^{l-1}v_q, v_1^{l-2}v_2^2, v_1^{l-2}v_2v_3, \dots, v_1^{l-2}v_2v_q, \dots, v_q^l]^T. \quad (37)$$

There exist matrices M_i and N_i satisfy

$$v^{[l]} = M_i v^{(l)}, \quad v^{(l)} = N_i v^{[l]}. \quad (38)$$

The nonlinear functions can be written as

$$\begin{aligned}f_i(x_i, u_i, v_i, \Delta_i) &= A_i(\Delta_i)x_i + B_i(\Delta_i)u_i + E_i(\Delta_i)v_i + \tilde{f}_i(x_i, u_i, v_i, \Delta_i) \\ h_i(x_i, u_i, v_i, \Delta_i) &= C_i(\Delta_i)x_i + D_i(\Delta_i)u_i + F_i(\Delta_i)v_i + \tilde{h}_i(x_i, u_i, v_i, \Delta_i)\end{aligned}\quad (39)$$

where

$$\begin{aligned}A_i(\Delta_i) &= \frac{\partial f_i}{\partial x_i}(0, 0, 0, \Delta_i), \quad B_i(\Delta_i) = \frac{\partial f_i}{\partial u_i}(0, 0, 0, \Delta_i) \\ C_i(\Delta_i) &= \frac{\partial h_i}{\partial x_i}(0, 0, 0, \Delta_i), \quad D_i(\Delta_i) = \frac{\partial h_i}{\partial u_i}(0, 0, 0, \Delta_i) \\ E_i(\Delta_i) &= \frac{\partial f_i}{\partial v_i}(0, 0, 0, \Delta_i), \quad F_i(\Delta_i) = \frac{\partial h_i}{\partial v_i}(0, 0, 0, \Delta_i).\end{aligned}$$

The smooth nonlinear residuals are \tilde{f}_i and \tilde{h}_i that vanish at $(x_i, u_i, v_i, \Delta_i) = (0, 0, 0, 0)$.

We design the control law for $t \in [t_i^k, t_i^{k+1})$ as

$$\begin{aligned}u_i &= k_i(x_i, z_i) = K_i^x x_i + K_i^z z_i \\ \dot{z}_i &= g_i(z_i, e_{v_i}) = \Sigma_1^i z_i + \Sigma_2^i e_{v_i} \\ e_{v_i} &= h_i(x_i, u_i, v_i, \Delta_i) - S_i v_i\end{aligned}\quad (40)$$

where z_i is the state of the dynamic compensator. The observation of v_i is $\hat{v}_i = \text{col}(\hat{\zeta}, \hat{f}_i)$, and (Σ_1^i, Σ_2^i) is the p-copy internal model pair of matrix A_i^k , which is defined as

$$\begin{aligned}A_i^k &= \begin{bmatrix} A_i^{[1]} & 0 & \dots & 0 \\ 0 & A_i^{[2]} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_i^{[k]} \end{bmatrix} \\ A_i^{[l]} &= M_i^l \left[\sum_{j=1}^l I_{q_i}^{(j-1)} \otimes A_i^v \otimes I_{q_i}^{(j-1)} \right] N_i^l\end{aligned}\quad (41)$$

where notations $A_i^{[l]}$ and $I_{q_i}^{(j-1)}$ are defined in (36) and (37) and q_i is the dimension of matrix A_i^v . Matrices M_i^l and N_i^l can be calculated based on (38).

The control laws K_i^x and K_i^z meet condition that

$$\tilde{A}_i^c(\Delta_i) = \begin{bmatrix} A_i(\Delta_i) + B_i(\Delta_i)K_i^x & B_i(\Delta_i)K_i^z \\ \Sigma_2^i(C_i(\Delta_i) + D_i(\Delta_i)K_i^x) & \Sigma_1^i + \Sigma_2^i D_i(\Delta_i)K_i^z \end{bmatrix} \quad (42)$$

is Hurwitz when $\Delta_i = 0$.

The second main result of this article on formation output control is as follows.

Theorem 2: Consider systems (1), (3), and (4) under Assumptions 1–5, and the solutions $x_i(v_i, \Delta_i)$ and $u_i(v_i, \Delta_i)$ of the regulation equations in (8) are degree $k > 0$ polynomials in v_i . Then, the nonlinear robust formation output control problem defined in Problem 1 is solved by dual adaptive observer (12) under event-triggered mechanism (15) and (16) and state feedback controller in (40), if conditions (13) and (42) are satisfied.

Proof: Substitution of state feedback law (40) into uncertain heterogeneous MASs (35) yields

$$\begin{aligned}\dot{\delta}_i &= \varepsilon_i(\delta_i, v_i, \Delta_i) \\ &= \tilde{A}_i^c(\Delta_i)\delta_i + \tilde{E}_i^c(\Delta_i)v_i + J_i^c(\Delta_i)\tilde{v}_i + \tilde{\varepsilon}_i(\delta_i, v_i, \Delta_i) \\ \dot{v}_i &= A_i^v v_i \\ e_i &= \rho_i(\delta_i, v_i, \Delta_i) \\ &= h_i(x_i, k_i(x_i, z_i), v_i, \Delta_i) - S_i v_i\end{aligned}\quad (43)$$

where $\delta_i = \text{col}(x_i, z_i)$, $\tilde{v}_i = \hat{v}_i - v_i$

$$\begin{aligned}\tilde{E}_i^c(\Delta_i) &= \begin{bmatrix} E_i(\Delta_i) \\ -\Sigma_2^i(F_i(\Delta_i) - S_i(\Delta_i)) \end{bmatrix} \\ J_i^c(\Delta_i) &= \begin{bmatrix} 0 \\ -\Sigma_2^i S_i(\Delta_i) \end{bmatrix} \\ \tilde{\varepsilon}_i(\delta_i, v_i, \Delta_i) &= \begin{bmatrix} \tilde{f}_i(x_i, k_i(x_i, z_i), v_i, \Delta_i) \\ \tilde{h}_i(x_i, k_i(x_i, z_i), v_i, \Delta_i) \end{bmatrix}\end{aligned}$$

and $\tilde{A}_i^c(\Delta_i)$ is given in (42). Based on [16, Th. 1], we have

$$\lim_{t \rightarrow \infty} \tilde{\zeta}_i = 0, \quad \lim_{t \rightarrow \infty} \tilde{v}_i = 0, \quad \lim_{t \rightarrow \infty} (e_i - e_{v_i}) = 0$$

where $\tilde{v}_i = \text{col}(0, \tilde{\zeta}_i, 0)$.

Let

$$\tilde{A}_i^c(0) = \begin{bmatrix} A_i(0) + B_i(0)K_i^x & B_i(0)K_i^z \\ \Sigma_2^i(C_i(0) + D_i(0)K_i^x) & \Sigma_1^i + \Sigma_2^i D_i(0)K_i^z \end{bmatrix} \quad (44)$$

be the nominal part of $\tilde{A}_i^c(\Delta_i)$ when $\Delta_i = 0$. Under Assumptions 1 and 3, the pair

$$\left(\begin{bmatrix} A_i(0) & 0 \\ \Sigma_2^i C_i(0) & \Sigma_1^i \end{bmatrix}, \begin{bmatrix} B_i(0) \\ \Sigma_2^i D_i(0) \end{bmatrix} \right) \quad (45)$$

is controllable. Hence, we can choose appropriate (K_i^x, K_i^z) such that matrices $\tilde{A}_i^c(0)$ for $i = 1, 2, \dots, n$ are Hurwitz. Based on Theorem 2.33 in [16], the matrix $\tilde{A}_i^c(\Delta_i)$ is also Hurwitz, and the system (35) at $(\delta_i, v_i) = (0, 0)$ is stable in the sense of Lyapunov.

According to the center manifold theorem in [16] and Assumptions 4 and 5, there exist a locally defined sufficiently smooth function $\bar{\delta}_i(v_i, \Delta_i)$ with $\bar{\delta}_i(0, 0) = 0$ that meets

$$\frac{\partial \bar{\delta}_i(v_i, \Delta_i)}{\partial v_i} A_i^v v_i = \varepsilon_i(\bar{\delta}_i(v_i, \Delta_i), v_i, \Delta_i) \quad (46)$$

for $i = 1, 2, \dots, n$. By partitioning

$$\bar{\delta}_i(v_i, \Delta_i) = \text{col}(\bar{x}_i(v_i, \Delta_i), \bar{z}_i(v_i, \Delta_i))$$

under control law (40), then (46) can be written as

$$\begin{aligned} \frac{\partial \bar{x}_i(v_i, \Delta_i)}{\partial v_i} A_i^v v_i &= f_i(\bar{x}_i(v_i, \Delta_i), \bar{u}_i(v_i, \Delta_i), v_i, \Delta_i) \\ \frac{\partial \bar{z}_i(v_i, \Delta_i)}{\partial v_i} A_i^v v_i &= \Sigma_1^i \bar{z}_i(v_i, \Delta_i) + \Sigma_2^i e_{v_i}(v_i, \Delta_i) \end{aligned} \quad (47)$$

where

$$\begin{aligned} \bar{u}_i(v_i, \Delta_i) &= K_i^x \bar{x}_i(v_i, \Delta_i) + K_i^z \bar{z}_i(v_i, \Delta_i) \\ e_{v_i}(v_i, \Delta_i) &= \rho_i(\bar{\delta}_i, v_i, \Delta_i). \end{aligned}$$

Since $\bar{x}_i(v_i, \Delta_i)$ and $\bar{u}_i(v_i, \Delta_i)$ are assumed to be degree k in v_i , then for any $k \geq 1$, functions $\bar{x}_i(v_i, \Delta_i)$, $\bar{z}_i(v_i, \Delta_i)$, and $e_{v_i}(v_i, \Delta_i)$ can be uniquely expressed as

$$\begin{aligned} \bar{x}_i(v_i, \Delta_i) &= \sum_{l=1}^k X_{il} v_i^{[l]} \\ \bar{z}_i(v_i, \Delta_i) &= \sum_{l=1}^k Z_{il} v_i^{[l]} \\ e_{v_i}(v_i, \Delta_i) &= \sum_{l=1}^k Y_{il} v_i^{[l]} \end{aligned} \quad (48)$$

where X_{il} , Z_{il} , and Y_{il} are constant matrices that may be related to Δ_i [16]. Substituting (48) into (47) and expanding (47) into power series, we can see

$$\begin{aligned} X_{il} A_i^{[l]} &= (A_i(\Delta_i) + B_i(\Delta_i) K_i^x) X_{il} \\ &\quad + B_i(\Delta_i) K_i^z Z_{il} + E_i(\Delta_i) \\ Z_{il} A_i^{[l]} &= \Sigma_1^i Z_{il} + \Sigma_2^i Y_{il} \\ Y_{il} &= (C_i(\Delta_i) + D_i(\Delta_i) K_i^x) X_{il} \\ &\quad + D_i(\Delta_i) K_i^z Z_{il} + F_{il} - S_{il} \end{aligned} \quad (49)$$

where (F_{il}, S_{il}) can be approximated to $[F_i(\Delta_i), S_i(\Delta_i)]$ based on Theorem 1 and (F_{il}, S_{il}) only depends on Z_{il} [16], [22].

Based on the p -copy internal model in Lemma 2 and k th-order output regulation in [16, Th. 5.7], (49) has a unique solution (X_{il}, Z_{il}) , which meets $Y_{il} = 0$. Defining

$$\begin{aligned} X_{cl} &= \text{col}(X_{1l}, Z_{1l}, X_{2l}, Z_{2l}, \dots, X_{nl}, Z_{nl}) \\ Y_{cl} &= \text{col}(Y_{1l}, Y_{2l}, \dots, Y_{nl}), \\ A^{[l]} &= \text{diag}(A_1^{[l]}, A_2^{[l]}, \dots, A_n^{[l]}) \\ A(\Delta_i) &= \text{diag}(\bar{A}_1^c(\Delta_1), \bar{A}_2^c(\Delta_1), \dots, \bar{A}_n^c(\Delta_n)) \\ E(\Delta_i) &= \text{diag}(\bar{E}_1^c(\Delta_1), \bar{E}_2^c(\Delta_1), \dots, \bar{E}_n^c(\Delta_n)) \\ C(\Delta_i) &= \text{diag}(\bar{C}_1^c(\Delta_1), \bar{C}_2^c(\Delta_1), \dots, \bar{C}_n^c(\Delta_n)) \\ F(\Delta_i) &= \text{diag}(\bar{F}_1^c(\Delta_1), \bar{F}_2^c(\Delta_1), \dots, \bar{F}_n^c(\Delta_n)) \end{aligned}$$

where $\bar{C}_i^c(\Delta_i) = [C_i(\Delta_i) + D_i(\Delta_i) K_i^x \quad D_i(\Delta_i) K_i^z]$ and $\bar{F}_i^c(\Delta_i) = F_i(\Delta_i) - S_i(\Delta_i)$. It implies

$$\begin{aligned} X_{cl} A^{[l]} &= A(\Delta_i) X_{cl} + E(\Delta_i) \\ Y_{cl} &= C(\Delta_i) X_{cl} + F(\Delta_i) = 0. \end{aligned} \quad (50)$$

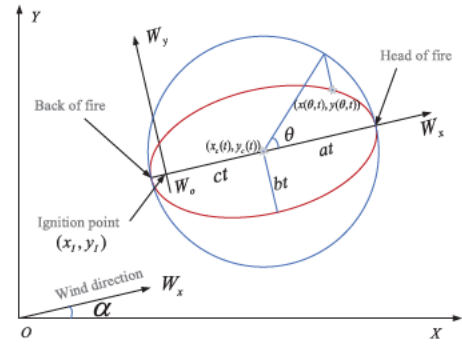


Fig. 2. Bushfires spread model.

From (50) and Theorem 1, we can see for $i = 1, 2, \dots, n$

$$\lim_{t \rightarrow \infty} e_{v_i} = 0, \quad \lim_{t \rightarrow \infty} e_i = 0, \quad \lim_{t \rightarrow \infty} y_i - y_0 - y_f = 0.$$

Therefore, the nonlinear robust formation problem is solved, which completes the proof. \square

Remark 6: Note that Theorems 1 and 2 provide a distributed robust solution to resolve the conflict between formation consensus requirements and unified nonlinear heterogeneous dynamics. The control law involves virtual leader observations and internal model information, instead of the states or outputs of neighbor agents. Hence, other agents can still work even when an agent is disrupted. The distributed controller can be extended to solve the consensus output problem in Definition 1 and formation control problem in Definition 2 if the formation dynamics and uncertainties/disturbances are not considered.

IV. VERIFICATION

In this section, an example of the reference formation design is given for tracking and patrolling tasks. Then, simulation and experiment are conducted to verify the control performance of the proposed robust strategy.

A. Reference Formation Design for Bushfire Tacking and Patrolling

The occurrence of multiple bushfires is unfortunately not uncommon in Australia. It has been observed that the shape of the fire is approximately an ovoid, which becomes more like an ellipse during the spread of the bushfires [33].

The ellipse model based on the main wind detection is shown in Fig. 2. The position of the bushfire range after time t from an ignition point is described as

$$\begin{aligned} x(\theta, t) &= ct + at \cos \theta \\ y(\theta, t) &= bt \sin \theta, \quad 0 \leq \theta < 2\pi \end{aligned} \quad (51)$$

where (W_x, W_o, W_y) is the wind coordinate system from the origin (ignition point) and (X, O, Y) denotes the inertial coordinate system. Coordinate (x_c, y_c) indicates the center of fire, and c is the moving rate of fire center in the direction of wind. Parameters a and b denote the linear fire spread rates along the wind direction and vertical wind directions, respectively. The angle of wind speed with respect to the X -axis of the inertial system is α . When $0 \leq \theta < 2\pi$,

the position set $[x(\theta, t), y(\theta, t)]$ represents the edge of the bushfires. The length of (at, bt) of the ellipse increases with time.

To perform the tracking task, a virtual leader is designed as the moving center of the ellipse. Define the state and output variables of the virtual leader as $x_0 = \text{col}(x_c, y_c, \dot{x}_c, \dot{y}_c)$ and $y_0 = \text{col}(x_c, y_c)$, respectively. Based on the description of model (3), the state and output matrices are obtained as

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (52)$$

In order to patrol the bushfire boundary, the TVF is designed with the purpose of making agents track the moving fire center and rotate around the edge of the ellipse. In the wind coordinate system (W_x, W_o, W_y) , a TVF based on the fire spread model in (51) is designed as follows:

$$\begin{aligned} f_i^x &= at \cos(\omega t + \theta_i^0) \\ f_i^y &= bt \sin(\omega t + \theta_i^0) \\ \dot{f}_i^x &= a \cos(\omega t + \theta_i^0) - a\omega \sin(\omega t + \theta_i^0) \\ \dot{f}_i^y &= b \sin(\omega t + \theta_i^0) + b\omega \cos(\omega t + \theta_i^0) \end{aligned} \quad (53)$$

where (f_i^x, f_i^y) represents the positions of the i th points, which are distributed over the ellipse, moving along the edge of the ellipse at an angular rate ω . The initial patrol angle is θ_i^0 .

Hence, the dynamics of TVF in (53) satisfy

$$\begin{bmatrix} \dot{f}_i^x \\ \dot{f}_i^y \\ \ddot{f}_i^x \\ \ddot{f}_i^y \end{bmatrix}_W = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \omega^2 & 0 & 0 & -\frac{2a\omega}{b} \\ 0 & \omega^2 & \frac{2b\omega}{a} & 0 \end{bmatrix} \begin{bmatrix} f_i^x \\ f_i^y \\ \dot{f}_i^x \\ \dot{f}_i^y \end{bmatrix}_W \quad (54)$$

where the subscript W denotes that the vector is in the wind coordinate system. Similarly, the subscript I represents that the vector is in the inertial coordinate system. The direction cosine matrices [34] between the two coordinate systems are

$$\text{DCM}_{W2I} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 & 0 \\ \sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & \cos \alpha & -\sin \alpha \\ 0 & 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

and $\text{DCM}_{I2W} = \text{DCM}_{W2I}^{-1} = \text{DCM}_{W2I}^T$.

If the state in (4) is taken as $f_i = [f_i^x, f_i^y, \dot{f}_i^x, \dot{f}_i^y]^T$, the system matrices in the inertial coordinate system become

$$A_i^f = \text{DCM}_{W2I} A_w^f \text{DCM}_{I2W} \\ A_w^f = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \omega^2 & 0 & 0 & -\frac{2a\omega}{b} \\ 0 & \omega^2 & \frac{2b\omega}{a} & 0 \end{bmatrix}, \quad C_i^f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (55)$$

Fig. 3 shows the tracking and patrolling movement of the reference formation model regarding the bushfire spread

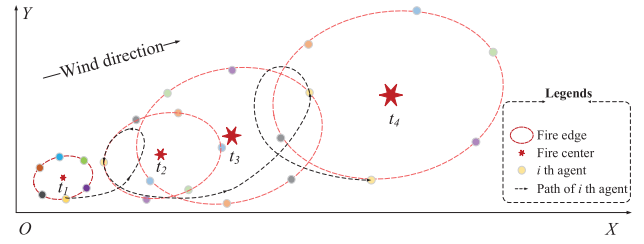


Fig. 3. Modeling of time-varying reference formation in bushfire edge tracking and patrolling tasks.

model. The fire center is regarded as the virtual leader, which moves at the speed c in the direction of the main wind. Meanwhile, the flaming range is gradually expanding as an ellipse. Under the virtual leader and TVF design, each agent moves around the fire center at an angular rate ω on the moving and gradually expanding ellipse. The path of one of the agents is also marked by a narrowed line at different times in Fig. 3. At each moment, multiple agents are distributed on the edge of the moving ellipse and patrolling around the edge.

B. Example 1—Simulation

An example is considered with a heterogeneous MAS consisting of two Qball2, two Qball-X4 UAVs [35], [36], and one Qbot2 UGV [10]. The nonlinear model of UAVs [36] is given as

$$\begin{aligned} \dot{\phi}_i^y &= p_i \\ \dot{p}_i &= k_i^m L_i u_{\phi_i} + (I_i^y - I_i^z) \dot{\theta}_i^x \dot{\psi}_i^z / I_i^x \\ \dot{\theta}_i^x &= q_i \\ \dot{q}_i &= k_i^m L_i u_{\theta_i} + (I_i^z - I_i^x) \dot{\phi}_i^y \dot{\psi}_i^z / I_i^y \\ \dot{\psi}_i^z &= r_i \\ \dot{r}_i &= k_i^m L_i u_{\psi_i} + (I_i^x - I_i^y) \dot{\phi}_i^y \dot{\theta}_i^x / I_i^z \\ \dot{p}_i^x &= v_i^x \\ \dot{v}_i^x &= (\sin \psi_i^z \sin \phi_i^y + \cos \psi_i^z \sin \theta_i^x \cos \phi_i^y) g + d_i^x \omega_x \\ \dot{p}_i^y &= v_i^y \\ \dot{v}_i^y &= (-\cos \psi_i^z \sin \phi_i^y + \sin \psi_i^z \sin \theta_i^x \cos \phi_i^y) g + d_i^y \omega_y \end{aligned} \quad (56)$$

where, for the i th agent, ϕ_i^y , θ_i^x , and ψ_i^z and p_i , q_i , and r_i are the Euler angles (roll, pitch, and yaw) and Euler angular rates, respectively. (p_i^x, p_i^y) and (v_i^x, v_i^y) represent the positions and velocities of the i th agent, respectively. The input variables and output variables of the UAV system are $u_i = [u_{\phi_i}, u_{\theta_i}, u_{\psi_i}]^T$ and $y_i = [p_i^x, p_i^y]^T$, respectively. Note that the dynamic in the z -direction is not considered here. The aircraft is assumed to fly in a plane with a safe altitude of 20 m for the tasks. Notations k_i^m and L_i indicate force coefficient and arm of force, respectively, and I_i^x , I_i^y , and I_i^z are moments of inertia in the X -, Y -, and Z -directions, respectively. $\omega = [\omega_x, \omega_y]$ is the disturbance, and (d_i^x, d_i^y) is the parameter indicating the influence of disturbance on the system.

The dynamic model of Qbot2 ground robot with disturbance $\omega = [\omega_x, \omega_y]$ can be modeled as

$$\begin{aligned} \dot{p}_i^x &= V_i \cos(\psi_i) \\ \dot{p}_i^y &= V_i \sin(\psi_i) \end{aligned}$$

$$\begin{aligned}\dot{V}_i &= a_i + \cos(\psi_i)d_i^x\omega_x + \sin(\psi_i)d_i^y\omega_y \\ \dot{\psi}_i &= r_i - \frac{1}{V_i}\sin(\psi_i)d_i^x\omega_x + \frac{1}{V_i}\cos(\psi_i)d_i^y\omega_y\end{aligned}\quad (57)$$

where (p_i^x, p_i^y) is the position of the i th UGV. Variables V_i and a_i represent the forward velocity and acceleration, respectively. The yaw angle and angular rate are ψ_i and r_i , respectively. Due to the nonholonomic constraint in (57) that the UGV system cannot slide, we can set

$$\begin{bmatrix} a_i \\ r_i \end{bmatrix} = \begin{bmatrix} \cos(\psi_i) & \sin(\psi_i) \\ -\frac{1}{V_i}\sin(\psi_i) & \frac{1}{V_i}\cos(\psi_i) \end{bmatrix} \begin{bmatrix} u_i^x \\ u_i^y \end{bmatrix}.\quad (58)$$

The system in (57) is converted to

$$\begin{aligned}\dot{p}_i^x &= v_i^x \\ \dot{p}_i^y &= v_i^y \\ \dot{v}_i^x &= u_i^x + d_i^x\omega_x \\ \dot{v}_i^y &= u_i^y + d_i^y\omega_y\end{aligned}\quad (59)$$

where v_i^x and v_i^y represent the velocities in the X - and Y -directions of the robot. The input variables and output variables of UGV system are $u_i = [u_i^x, u_i^y]^T$ and $y_i = [p_i^x, p_i^y]^T$, respectively.

To perform the tracking and patrolling tasks, heterogeneous MASs with different dynamic and orders are considered, where the first four agents ($i = 1, 2, 3,$ and 4) are UAVs for patrolling around the moving and expanding brushfire edge and agent 5 is a UGV for tracking at the back of fire. For $i = 1$ and 2 , $k_i^m = 12 \text{ N} + \Delta k_i^m$, $L_i = 0.2 \text{ m} + \Delta L_i$, $I_i^x = I_i^y = 0.03 \text{ kg} \cdot \text{m}^2$, and $I_i^z = 0.04 \text{ kg} \cdot \text{m}^2$, whereas for $i = 3$ and 4 , $k_i^m = 120 \text{ N} + \Delta k_i^m$, $L_i = 0.2 \text{ m} + \Delta L_i$, $I_i^x = I_i^y = 0.03 \text{ kg} \cdot \text{m}^2$, and $I_i^z = 0.04 \text{ kg} \cdot \text{m}^2$, where the system uncertainties are given as $\Delta k_{1,2,3,4}^m = (0.1, 0, -0.1, -0.1) \text{ N}$ and $\Delta L_{1,2,3,4} = (0.02, 0.01, 0, -0.05) \text{ m}$. The disturbances are assumed to be the main wind disturbances with $A_w = \mathbf{0}$, $w_0 = [w_0^x, w_0^y]^T = [2, 1]$ in (2), and $d_i^x = 0.1, d_i^y = 0.1$ when $i = 1, 2, 3,$ and 4 , while $d_5^x = 0.15$ and $d_5^y = 0.15$. The dynamics of fire center regarding as virtue leader are given in (52) with the initial state $x_0 = [0 \ 1 \ 2 \ 1]^T$. The parameters in the design of TVF in (55) are chosen as $c = [2 \ 1]^T$, $a = 0.2$, $b = 0.1$, $\omega = 0.8$, and $\alpha = 26.5^\circ$. For the TVF of UGV, patrolling angular rate $\omega = 0$ is to only track the back of fire. The initial patrolling angles of TVF are $\theta^0 = (0, \pi/2, \pi, 3\pi/2, 0)$. The communication topology of MAS with virtual leader is shown in Fig. 4(a). The parameters of distributed event-triggered observer (12) are set as $R = 2I$ and $Q = 0.1I$. Based on Theorems 1 and 2, the simulation results of observation errors and formation control are shown in Figs. 5 and 6 and Figs. 7 and 8, respectively.

As we can see from Fig. 5, the errors of the distributed adaptive observer for the fire center and the main wind disturbances converge to zero within 15 s. Meanwhile, the event-triggered instants of each agent are given in Fig. 6, where the average trigger interval is 0.4099 s. The positions of agents are drawn every 5 s in Fig. 7. The trajectory of UAV1 and UGV1 is also plotted. It shows that the movements of agents match the desired reference formation for bushfire edge tracking and patrolling. Based on the adaptive observer in (12) and robust

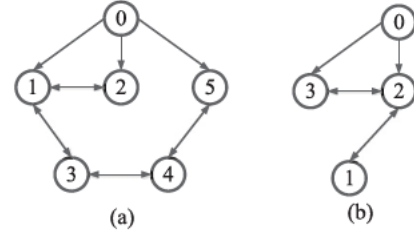


Fig. 4. Communication topologies of heterogeneous MASs. (a) Topology in Example 1. (b) Topology in Example 2.

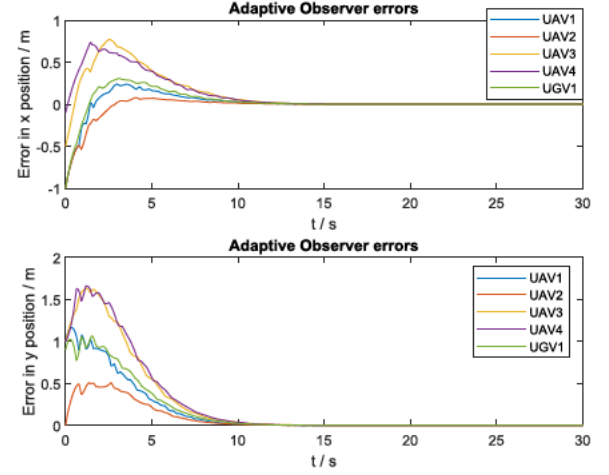


Fig. 5. Observation errors for the virtual leader (fire center) and wind disturbances.

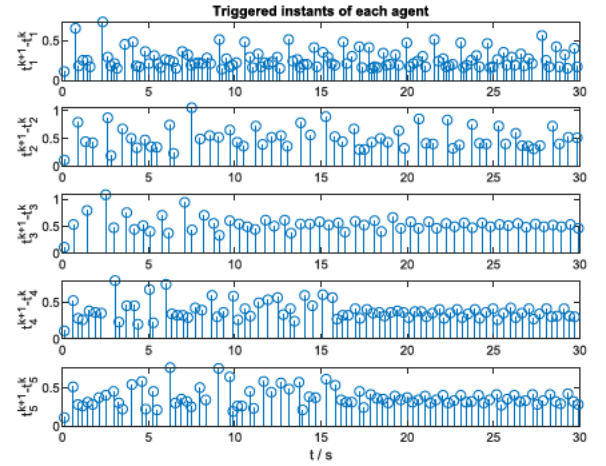


Fig. 6. Event-triggered instants of each agent.

formation controllers presented in (40), the formation errors of nonlinear heterogeneous MAS (1) with uncertainties and disturbances converge to zero within 20 s, as shown in Fig. 8.

C. Example 2—Experiment

For experimental verification, we give a simple example involving only UGVs. As shown in Fig. 9, one of the four Qbot2 is the fire center motion simulator; the other three are mobile land vehicles labeled as UGV1, UGV2, and UGV3. Each of the UGVs has equipped with a Gumstix DuoVero Zephyr onboard computer with an integrated 802.11 b/g/n

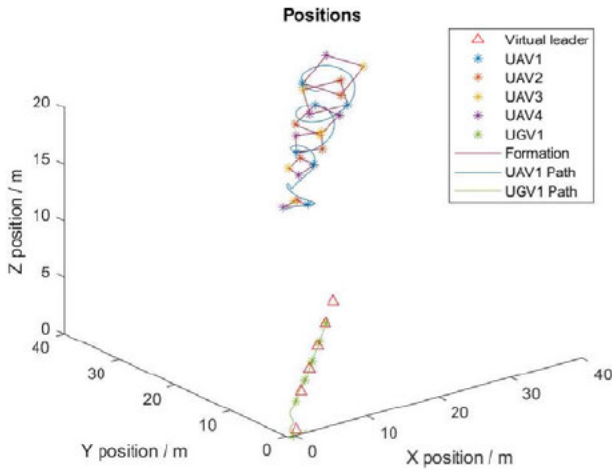


Fig. 7. Positions of heterogeneous MASs under system uncertainties and disturbances with event-triggered strategy.

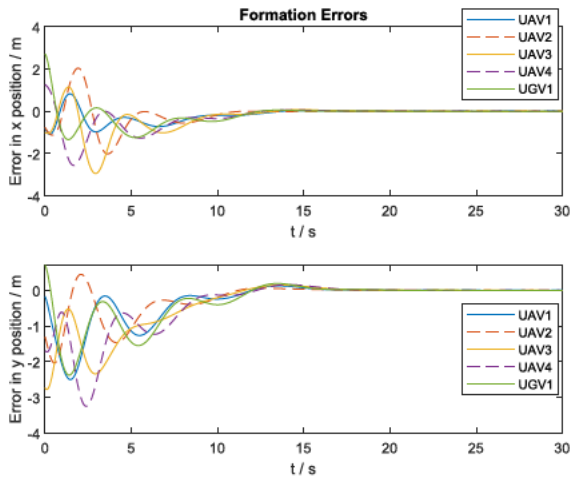


Fig. 8. Formation errors of heterogeneous MASs under system uncertainties and disturbances.

Wi-Fi module, while only UGV2 and UGV3 have onboard Kinect RGBD depth cameras. The global OptiTrack system with eight cameras and the onboard cameras of two Qbot2 are used for detection and positioning. They can communicate with each other by a local wireless network via TCP/IP protocol. There is a control station with QUARC real-time control software and Motive real-time software to monitor the status of UGVs wirelessly. Their communication topology is shown in Fig. 4(b).

Based on the size of the experimental site, we design the parameters for TVF as $C = [0.1 \ 0]^T$, $a = b = 0.001$, $\omega = 0.1$, $\alpha = 0$, and $\theta^0 = (0, 2\pi/3, 4\pi/3)$. For disturbances, $\omega_0 = [0.1, 0]^T$, and the disturbance coefficients in the UGV system (57) are $d_{1,2,3}^x = (1, 1, 0.1)$ and $d_{1,2,3}^y = (0.1, 0.1, 1)$. Observers and formation controllers are designed based on Theorems 1 and 2 under $R = 2I$ and $Q = 0.1I$ in (12). The experimental results are shown in Figs. 9–16.

From Fig. 10, the error under the dual adaptive observer converges to almost zero in 20 s, where the event-triggered interval is shown in Fig. 11. There are some observer errors due to the time delay. Because agent 2 is connected to all other

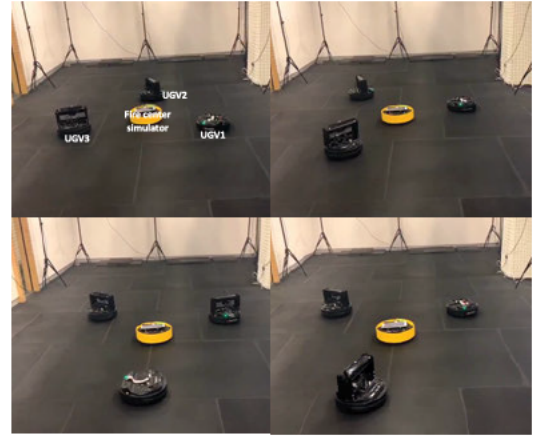


Fig. 9. Experimental photographs of the TVF.

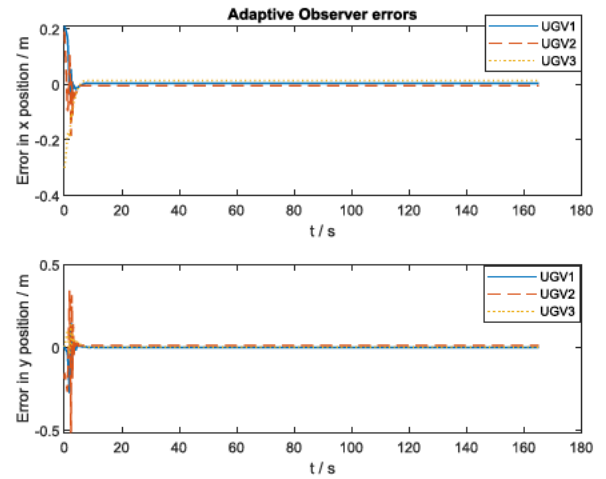


Fig. 10. Observation errors for virtual leader and disturbances.

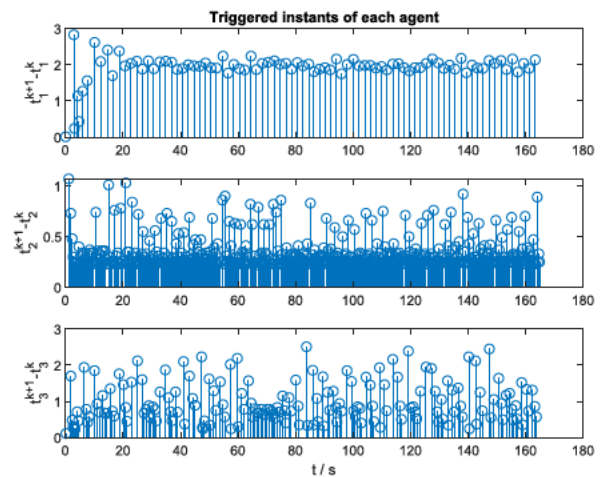


Fig. 11. Event-triggered instants of each agent.

nodes in their communication topology, it is reasonable that the interaction frequency of UGV2 is the highest in Fig. 11 and the oscillation is large before 20 s in Fig. 10. Note that the default interaction interval of the experimental platform

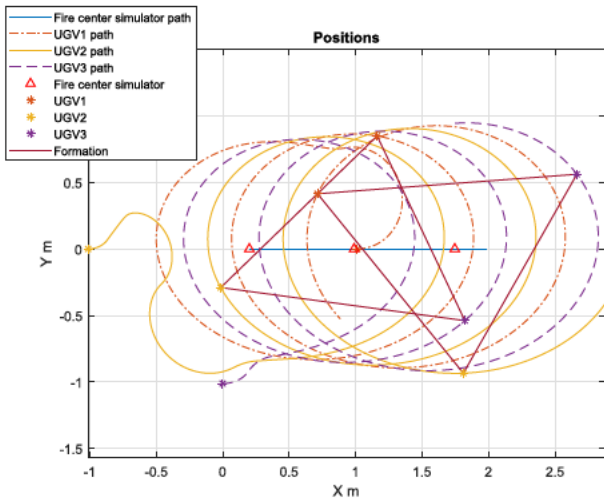


Fig. 12. Positions of agents.

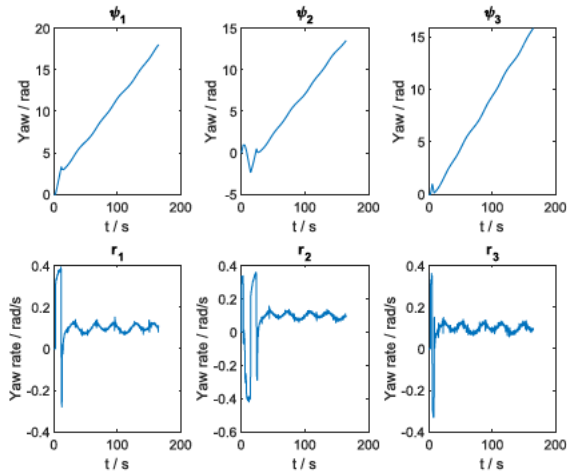


Fig. 13. Yaw and yaw rate of agents.

is 0.01 s in the absence of the event-triggered strategy. After adding the dual adaptive event-triggered strategy, the average interval time of the entire MAS is 0.6578 s, which means that the communication frequency is greatly reduced by more than 60 times, and the propagation rate of cyber-failures or attacks will also be significantly slowed down. Therefore, the proposed strategy reduces the frequency of interaction while ensuring a desired system performance.

The trajectories of UGVs are shown in Figs. 9 and 12, where three UGVs patrol and track around the center of the fire simulator, on a gradually expanding ellipse. Comparing Fig. 3 with Fig. 12, the movements of agents match the desired reference formation for the tasks. Figs. 14–16 show the changes of yaw angles, yaw angular rates, velocities, and formation tracking trajectories of agents. Compared with UGV1 and UGV3, the states of UGV2 also have larger oscillations before 40 s due to the results of the event-triggered observer. From the formation tracking results, the positions of three UGVs can almost track the formation reference command within 40 s. Note that there are errors in formation control due to the uncertainty of the experimental environment and time delay.

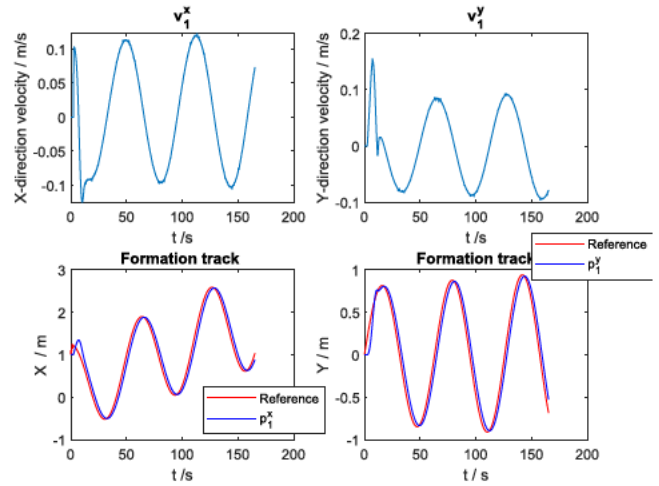


Fig. 14. Velocity and formation reference tracking of UGV1.

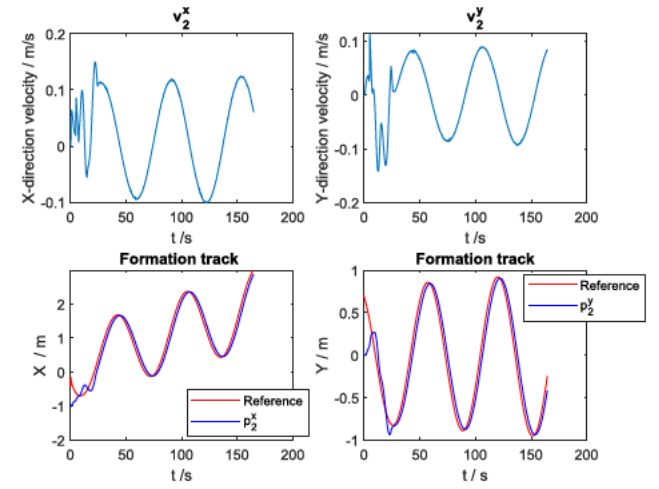


Fig. 15. Velocity and formation reference tracking of UGV2.

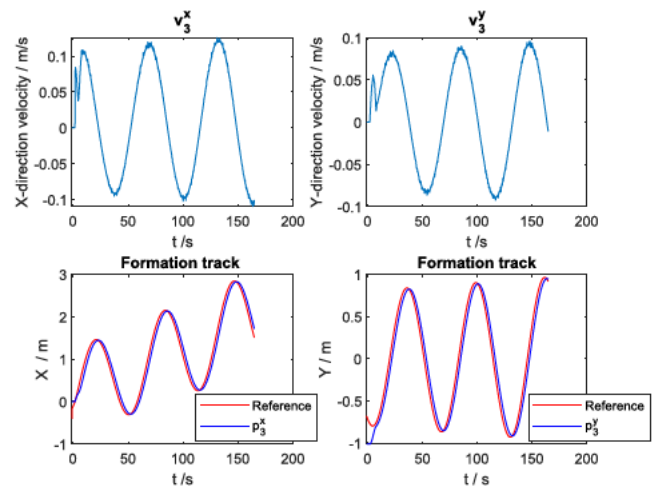


Fig. 16. Velocity and formation reference tracking of UGV3.

V. CONCLUSION

In this article, a solution to the nonlinear robust formation output control problem is provided for unified

heterogeneous uncertain MASs. First, a distributed dual adaptive event-triggered observer is developed to estimate the states of the virtual leader and disturbances that rely only on local information. Then, a robust formation controller is designed based on a nonlinear internal model compensator to tolerate uncertainties and disturbances. Simulation and experimental examples demonstrate that the proposed control strategies are effective in providing a theoretical reference for the application of formation control in cooperative tracking and patrolling tasks in search and rescue operations. There are a number of challenges remaining and deserving our future investigation, including fault-tolerant problems and performance optimization for heterogeneous MASs. A possible solution is to combine existing methods, such as event-based fault-tolerant control [26], switched frameworks [5], [37], fuzzy logic [38], neural network [39], and reinforcement learning [40] with our robust formation control strategy.

REFERENCES

- [1] J. Qin, Q. Ma, Y. Shi, and L. Wang, "Recent advances in consensus of multi-agent systems: A brief survey," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4972–4983, Jun. 2016.
- [2] P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 161–175, Jan. 2021.
- [3] B. Wang, J. Wang, B. Zhang, W. Chen, and Z. Zhang, "Leader-follower consensus of multivehicle wirelessly networked uncertain systems subject to nonlinear dynamics and actuator fault," *IEEE Trans. Automat. Sci. Eng.*, vol. 15, no. 2, pp. 492–505, Apr. 2017.
- [4] Z.-G. Wu, Y. Xu, Y.-J. Pan, H. Su, and Y. Tang, "Event-triggered control for consensus problem in multi-agent systems with quantized relative state measurements and external disturbance," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 7, pp. 2232–2242, Jul. 2018.
- [5] J. Wang, Y. Wang, H. Yan, J. Cao, and H. Shen, "Hybrid event-based leader-following consensus of nonlinear multiagent systems with semi-Markov jump parameters," *IEEE Syst. J.*, early access, Oct. 23, 2020, doi: [10.1109/JSYST.2020.3029156](https://doi.org/10.1109/JSYST.2020.3029156).
- [6] X. Dong, Y. Hua, Y. Zhou, Z. Ren, and Y. Zhong, "Theory and experiment on formation-containment control of multiple multirotor unmanned aerial vehicle systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 1, pp. 229–240, Jan. 2018.
- [7] B.-S. Chen, C.-P. Wang, and M.-Y. Lee, "Stochastic robust team tracking control of multi-UAV networked system under Wiener and Poisson random fluctuations," *IEEE Trans. Cybern.*, early access, Jan. 10, 2020, doi: [10.1109/TCYB.2019.2960104](https://doi.org/10.1109/TCYB.2019.2960104).
- [8] C.-C. Yang and T.-H. Cheng, "Leader-follower cooperative swinging by UAVs without interagent communication," *Int. J. Control*, Jan. 2020, doi: [10.1080/00207179.2020.1779959](https://doi.org/10.1080/00207179.2020.1779959).
- [9] J. Fischer, C. Lieberoth-Leden, J. Fottner, and B. Vogel-Heuser, "Design, application, and evaluation of a multiagent system in the logistics domain," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1283–1296, Jul. 2020, doi: [10.1109/TASE.2020.2979137](https://doi.org/10.1109/TASE.2020.2979137).
- [10] B. Yan, C. Wu, and P. Shi, "Formation consensus for discrete-time heterogeneous multi-agent systems with link failures and actuator/sensor faults," *J. Franklin Inst.*, vol. 356, no. 12, pp. 6547–6570, Aug. 2019.
- [11] Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and Y. Shen, "Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control," *Automatica*, vol. 93, pp. 26–32, Jul. 2018.
- [12] W. He, B. Yan, and C. Wu, "Distributed cooperative formation control for multi-agent systems based on robust adaptive strategy," *ICIC Exp. Lett.*, vol. 14, no. 7, pp. 661–668, 2020, doi: [10.24507/icicel.14.07.661](https://doi.org/10.24507/icicel.14.07.661).
- [13] L. Ji, T. Gao, and X. Liao, "Couple-group consensus for cooperative-competitive heterogeneous multiagent systems: Hybrid adaptive and pinning methods," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access, Nov. 15, 2019, doi: [10.1109/TSMC.2019.2951787](https://doi.org/10.1109/TSMC.2019.2951787).
- [14] B. Yan, P. Shi, C.-C. Lim, C. Wu, and Z. Shi, "Optimally distributed formation control with obstacle avoidance for mixed-order multi-agent systems under switching topologies," *IET Control Theory Appl.*, vol. 12, no. 13, pp. 1853–1863, 2018.
- [15] X. Li and P. Shi, "Cooperative fault-tolerant tracking control of heterogeneous hybrid-order mechanical systems with actuator and amplifier faults," *Nonlinear Dyn.*, vol. 98, no. 1, pp. 447–462, Oct. 2019.
- [16] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia, PA, USA: SIAM, 2004.
- [17] Y. Su and J. Huang, "Cooperative robust output regulation of a class of heterogeneous linear uncertain multi-agent systems," *Int. J. Robust Nonlinear Control*, vol. 24, no. 17, pp. 2819–2839, Nov. 2014.
- [18] Y. Xu, M. Fang, Y. Pan, K. Shi, and Z. Wu, "Event-triggered output synchronization for nonhomogeneous agent systems with periodic denial-of-service attacks," *Int. J. Robust Nonlinear Control*, vol. 31, no. 6, pp. 1851–1865, Apr. 2021, doi: [10.1002/rnc.5223](https://doi.org/10.1002/rnc.5223).
- [19] S. Dong, G. Chen, M. Liu, and Z.-G. Wu, "Cooperative neural-adaptive fault-tolerant output regulation for heterogeneous nonlinear uncertain multiagent systems with disturbance," *Sci. China Inf. Sci.*, vol. 64, May 2021, Art. no. 172212, doi: [10.1007/s11432-020-3122-6](https://doi.org/10.1007/s11432-020-3122-6).
- [20] H. Du, G. Wen, D. Wu, Y. Cheng, and J. Lü, "Distributed fixed-time consensus for nonlinear heterogeneous multi-agent systems," *Automatica*, vol. 113, Mar. 2020, Art. no. 108797.
- [21] J. R. Klotz, T.-H. Cheng, and W. E. Dixon, "Robust containment control in a leader-follower network of uncertain Euler-Lagrange systems," *Int. J. Robust Nonlinear Control*, vol. 26, no. 17, pp. 3791–3805, Nov. 2016.
- [22] S. Li, J. Zhang, X. Li, F. Wang, X. Luo, and X. Guan, "Formation control of heterogeneous discrete-time nonlinear multi-agent systems with uncertainties," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4730–4740, Jun. 2017.
- [23] W. Jiang, G. Wen, Z. Peng, T. Huang, and R. Rahmani, "Fully distributed formation-containment control of heterogeneous linear multi-agent systems," *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3889–3896, Sep. 2018.
- [24] H. Cai, F. L. Lewis, G. Hu, and J. Huang, "The adaptive distributed observer approach to the cooperative output regulation of linear multi-agent systems," *Automatica*, vol. 75, pp. 299–305, Jan. 2017.
- [25] R. Yang, H. Zhang, G. Feng, H. Yan, and Z. Wang, "Robust cooperative output regulation of multi-agent systems via adaptive event-triggered control," *Automatica*, vol. 102, no. 6, pp. 129–136, Apr. 2019.
- [26] Y. Xu and Z.-G. Wu, "Distributed adaptive event-triggered fault-tolerant synchronization for multiagent systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 2, pp. 1537–1547, Feb. 2020.
- [27] Z. Li, M. Chen, and Z. Ding, "Distributed adaptive controllers for cooperative output regulation of heterogeneous agents over directed graphs," *Automatica*, vol. 68, no. 6, pp. 179–183, Jun. 2016.
- [28] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 35, no. 2, pp. 131–140, Feb. 1990.
- [29] L. Pajonon, "The role of exosystems in output regulation," *IEEE Trans. Autom. Control*, vol. 59, no. 8, pp. 2301–2305, Aug. 2014.
- [30] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. London, U.K.: Springer-Verlag, 2009.
- [31] X.-W. Bu, X.-Y. Wu, Y.-X. Chen, and R.-Y. Bai, "Design of a class of new nonlinear disturbance observers based on tracking differentiators for uncertain dynamic systems," *Int. J. Control Automat. Syst.*, vol. 13, no. 3, pp. 595–602, 2015.
- [32] X. Bu, X. Wu, R. Zhang, Z. Ma, and J. Huang, "Tracking differentiator design for the robust backstepping control of a flexible air-breathing hypersonic vehicle," *J. Franklin Inst.*, vol. 352, no. 4, pp. 1739–1765, Apr. 2015.
- [33] G. L. W. Perry, "Current approaches to modelling the spread of wildland fire: A review," *Prog. Phys. Geography, Earth Environ.*, vol. 22, no. 2, pp. 222–245, Jun. 1998.
- [34] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of Unmanned Aerial Vehicles*. Dordrecht, The Netherlands: Springer, 2015.
- [35] M. Y. Xu, P. Yang, Y. X. Wang, and Q. B. Shu, "Observer-based multi-agent system fault upper bound estimation and fault-tolerant consensus control," *Int. J. Innov. Comput. Inf. Control*, vol. 15, no. 2, pp. 519–534, 2019.
- [36] Z. Liu, C. Yuan, Y. Zhang, and J. Luo, "A learning-based fault tolerant tracking control of an unmanned quadrotor helicopter," *J. Intell. Robotic Syst.*, vol. 84, nos. 1–4, pp. 145–162, Dec. 2016.
- [37] Y. Wang, X. Hu, K. Shi, X. Song, and H. Shen, "Network-based passive estimation for switched complex dynamical networks under persistent dwell-time with limited signals," *J. Franklin Inst.*, vol. 357, no. 15, pp. 10921–10936, Oct. 2020.
- [38] X. Bu and Q. Qi, "Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design," *IEEE Trans. Fuzzy Syst.*, early access, Nov. 11, 2020, doi: [10.1109/TFUZZ.2020.3036706](https://doi.org/10.1109/TFUZZ.2020.3036706).

- [39] X. Bu, "Air-breathing hypersonic vehicles funnel control using neural approximation of non-affine dynamics," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 5, pp. 2099–2108, Oct. 2018.
- [40] R. Watanuki, T. Horiuchi, and T. Aodal, "Vision-based behavior acquisition by deep reinforcement learning in multi-robot environment," *ICIC Exp. Lett. B, Appl.*, vol. 11, no. 3, pp. 237–244, 2020.



Bing Yan received the B.Sc. degree in automation and the M.Sc. degree in control theory and control engineering from Northwestern Polytechnical University, Xi'an, China, in 2012 and 2015, respectively. She is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA, Australia.

Her research interests include flight control, formation control, and multiagent systems.



Peng Shi (Fellow, IEEE) received the Ph.D. degree in electrical engineering from The University of Newcastle, Callaghan, NSW, Australia, in 1994, the Doctor of Science degree from the University of Glamorgan, Pontypridd, U.K., in 2006, and the Doctor of Engineering degree from The University of Adelaide, Adelaide, SA, Australia, in 2015.

He is currently a Professor at The University of Adelaide. His research interests include system and control theory and applications to autonomous and robotic systems, network systems, and cyber-physical systems.

Dr. Shi is a fellow of the Institution of Engineering and Technology and the Institute of Engineers, Australia. He has served on the Editorial Board of a number of journals, including *Automatica*, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, *IEEE TRANSACTIONS ON CYBERNETICS*, *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, and *IEEE CONTROL SYSTEMS LETTERS*. He is also the President of Systems and Cybernetic Science and the Vice President and Distinguished Lecturer of the IEEE Systems, Man, and Cybernetics Society.



Cheng-Chew Lim (Life Senior Member, IEEE) received the B.Sc. degree (Hons.) in electronic and electrical engineering and the Ph.D. degree in electronic and electrical engineering from Loughborough University, Leicestershire, U.K., in 1977 and 1981, respectively.

He is currently a Professor at The University of Adelaide, Adelaide, SA, Australia. His research interests include control and systems theory, autonomous systems, machine learning, and optimization techniques and applications.

Dr. Lim has served as an Editorial Board Member for the *IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS* and *Journal of Industrial and Management Optimization*.

Chapter 5

Optimal Robust Formation Control for Heterogeneous Multi-agent Systems Based on Reinforcement Learning

Statement of Authorship

Title of Paper	Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	B. Yan, P. Shi, C. -C. Lim and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," International Journal of Robust and Nonlinear Control, vol. 32, no. 5, pp. 2683–2704, 2022.

Principal Author

Name of Principal Author (Candidate)	Bing Yan		
Contribution to the Paper	Conceptualization, methodology, validation and writing-original draft		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/07/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Peng Shi		
Contribution to the Paper	Polishing, checking and verification		
Signature		Date	25 July 2022

Name of Co-Author	Cheng-Chew Lim		
Contribution to the Paper	Review, refine and validate		
Signature		Date	26/7/22

Name of Co-Author	Zhiyuan Shi		
Contribution to the Paper	Programming and writing-review		
Signature		Date	26 Jul 2022


5.1 Introduction

Considering unknown heterogeneous multi-agent systems (MAS) and an unknown exosystem, this chapter provides a novel reinforcement learning (RL)-based distributed formation optimization to achieve time-varying formation (TVF) without collisions. Three new off-policy RL algorithms are proposed to learn the optimal policies of each agent in real time. An observed model-based RL algorithm or a model-free RL algorithm can be used to estimate the dynamics and states of a reference exosystem. Another model-free RL algorithm is integrated with a collision-free formation controller to solve TVF optimization problems. Compared with most existing approaches focusing on quadratic objective functions, the developed control method addresses the non-quadratic optimization problem when the system model is completely unknown. Comparative simulations demonstrate the real-time learning performance and dynamic collision avoidance capability of a UAV-UGV heterogeneous MAS.

5.2 Publication

B. Yan, P. Shi, C. -C. Lim and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 5, pp. 2683–2704, 2022.

Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning

Bing Yan¹ | Peng Shi¹  | Cheng-Chew Lim¹ | Zhiyuan Shi²

¹School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, South Australia, Australia

²School of Automation, Northwestern Polytechnical University, Xi'an, P.R. China

Correspondence

Peng Shi, School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia.
Email: peng.shi@adelaide.edu.au

Funding information

Australian Research Council,
Grant/Award Number: DP170102644

Abstract

In this article, a reinforcement learning (RL)-based robust control strategy is proposed for uncertain heterogeneous multi-agent systems to achieve optimal collision-free time-varying formations. Without using any global information, a fully distributed adaptive observer is developed to estimate both dynamics and states of the reference and disturbance systems. The observer parameters are found by an observed model-based or a model-free off-policy RL algorithm. Using the internal model principle, a novel optimal robust formation control strategy is developed based on another proposed off-policy RL algorithm. The algorithm addresses the nonquadratic optimization problem when the system model is completely unknown. Taking the bushfire edge tracking and patrolling task for an unmanned aerial vehicle-unmanned ground vehicle heterogeneous system as an example, the effectiveness and robustness of the developed control strategy are verified by simulations.

KEYWORDS

adaptive observer, heterogeneous multi-agent systems, reinforcement learning, robust formation control

1 | INTRODUCTION

Driven by the development of artificial intelligence and computer science, multi-agent collaborative intelligence has been extensively studied in recent decades.¹⁻³ As a typical multi-agent collaboration technology, formation control aims to drive intractable agents to move as desired geometric shapes for better adapting to the tasks and environment.⁴⁻⁶ The time-varying formation (TVF) has more advantages to deal with environmental changes,^{6,7} including collision avoidance and dynamical target tracking. However, the formation technology of large-scale multi-agents adapt to practical tasks remains a challenge due to the constraints in the heterogeneity of the system, the locality of interactive information, the unknown and uncertainties of the system model, and the complexity of the environment.

In practical multi-agent applications, there are almost no completely homogeneous agents with the same individual parameters and the clock synchronization capability. Heterogeneous systems can involve different types of agents, such as unmanned aerial vehicles-unmanned ground vehicles (UAVs-UGVs), and are required to be more flexible in allocating tasks compared with homogeneous agents. One of the basic heterogeneous systems, mixed-order MASs were studied on consensus problems⁸ and formation control.^{9,10} For unified heterogeneous MASs with different state orders and dynamics, their common interest variables can be regarded as the outputs with the same

Abbreviations: HJB, Hamilton–Jacobi–Bellman; MAS, multi-agent system; RL, reinforcement learning; TVF, time-varying formation; UAV, unmanned aerial vehicle; UGV, unmanned ground vehicle.

dimensions. Therefore, the output consensus is more reasonable than the state consensus.¹¹ Inspired by output regulation problems for single systems to track a reference exosystem,¹² the output consensus problems were studied for heterogeneous MASs.^{13,14} The framework is extended to solve the formation control problems of unified heterogeneous systems.⁵

Formation control can be categorized into the centralized methods^{15,16} and the distributed methods^{17,18} based on whether the agent relies on global information or local information. A great number of distributed methods have been proposed to offer more flexibility and robustness. However, some existing distributed methods^{19,20} also involve prior global communication knowledge. In order to overcome the drawbacks, adaptive observers have been introduced to estimate the common references and decouple the system dynamics from communication networks.^{13,21} Although the restrictions of communication knowledge are removed, the proposed observers are still based on the known models of both reference and agent itself in most cases. To estimate the model of exosystem, an observer has been proposed to solve output regulation problems recently,²² but the knowledge of minimum eigenvalues of the communication Laplacian matrix is needed in the control law design. It is technically challenging to remove both topology and model information in fully distributed formation control for unified uncertain heterogeneous MASs.

Furthermore, since the energy of agents is limited and collisions may occur in dynamic environments, optimal control with collision avoidance is another requirement that must be considered in formation control. While considerable advancements have been made for dynamic path planning using such approaches as heuristic²³ and artificial intelligence.²⁴ Most algorithms are still model-based, the optimal control design is still a challenge for agents to achieve collision-free formation with optimization under unknown dynamics.

Therefore, it is necessary to develop optimal robust formation control technologies to deal with practical constraints arising from uncertainties, disturbances, and unknown models. The robust internal model control was proposed to achieve output consensus under uncertainties and disturbance without solving the output regulation equation.^{13,25} Taking the reference formation into consideration, this method has been extended to ensure the formation control stability of closed-loop heterogeneous systems.²⁶ In order to solve the system optimization problems, mature algorithms based on the Hamilton–Jacobi–Bellman (HJB) equation are widely used when the system model is known.²⁷

In recent years, reinforcement learning (RL) algorithms have gained popularity in industrial communities to solve optimization problems with unknown models.^{28–30} RL is one of the machine learning algorithms that learn the optimal strategies by observing the responses to the environment.^{31,32} The approaches used for RL are mostly based on on-policy RL and off-policy RL.³³ The target policy and behavior policy are the same in the former and different in the latter. The first on-policy RL based on adaptive dynamic programming³⁴ only required a partial understanding of system dynamics. By contrast, off-policy RL with different policies has more advantages in mining data and stimulating the system.^{35,36} Therefore, off-policy RL has been more widely deployed to solve the problem of optimal control with unknown models in MASs. Note that in most existing off-policy RL algorithms, only quadratic objective functions are considered.^{36–38} The functions to describe collision avoidance cost usually cannot be written in quadratic forms. For large-scale multi-agent systems, the issues about heterogeneity and unknowns of the system, the coupling of data interaction, and the uncertain obstacles in the environment, all can pose challenges to the optimization of data-driven formation control and dynamic obstacle avoidance.

Therefore, a learning-based robust formation control strategy under mentioned constraints is investigated in the article. The main contributions in this article are as follows:

1. Design a TVF control strategy for a unified unknown heterogeneous MAS with different orders and different dynamics to adapt to tasks such as tracking and patrolling.
2. Propose an adaptive observer with observed model-based and model-free iterative RL algorithms to estimate the dynamics and states of a reference system consisting of a virtual leader and disturbance without requiring any global information.
3. Develop a robust formation output controller to achieve the optimal collision-free heterogeneous formation control based on off-policy RL, which is capable of solving the nonquadratic optimization problem without using any model information.

The notation used in this article is standard. X^{-1} and Y^T represent the inverse of nonsingular and square matrix X and transpose of matrix Y , respectively. $\|X\|$ is the norm of matrix X . Notation \otimes denotes the Kronecker product, and blkdiag represents the block diagonal concatenation of matrix input arguments. $\text{col}(x_1, x_2, \dots, x_n)$ is the column vector consisting of vectors (x_1, x_2, \dots, x_n) . Function $\text{vec}(X)$ reshape the elements in matrix X into a column vector.

2 | PROBLEM FORMULATION

A unified heterogeneous uncertain MAS with n agents of different orders can be modeled as

$$\dot{x}_i = \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{D}_i \omega, \quad y_i = \tilde{C}_i x_i, \quad i = 1, 2, \dots, n, \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}$, and $y_i \in \mathbb{R}^p$ represent the state, input, and output variables of the i th agent, respectively. Matrices

$$\tilde{A}_i = A_i + \Delta A_i, \quad \tilde{B}_i = B_i + \Delta B_i, \quad \tilde{C}_i = C_i + \Delta C_i, \quad \tilde{D}_i = D_i + \Delta D_i$$

are the system state, input, output, and disturbance matrices with uncertainties $\Delta \star_i$, where $\star = A, B, C, D$. The uncertainty matrices are assumed to be bounded, and belong to an open neighborhood W of the original point.

The dynamics of disturbance signal $\omega \in \mathbb{R}^{n_\omega}$ is described by

$$\dot{\omega} = A_\omega \omega. \tag{2}$$

Generally, the nonlinear function of disturbances can be expanded by the Fourier series and approximated as a combined signal via trigonometric functions. The state matrix A_ω can always be found from the trigonometric functions to predict changes of disturbance signals.

To implement the fully distributed formation control for heterogeneous MASs, a virtual leader is designed in the desired formation as

$$\dot{x}_0 = A_0 x_0, \quad y_0 = C_0 x_0, \tag{3}$$

where $x_0 \in \mathbb{R}^{n_0}$ and $y_0 \in \mathbb{R}^p$ are state and output variables of the virtual leader, respectively. The state matrix and output matrix of the system are A_0 and C_0 , respectively. Note that only some agents get the information from the virtual leader.

In addition to a virtual leader, the shape of formation is illustrated by the TVF system

$$\dot{f}_i = A_i^f f_i, \quad y_i^f = C_i^f f_i, \quad i = 1, 2, \dots, n, \tag{4}$$

where $f_i \in \mathbb{R}^{n_f}$ and $y_i^f \in \mathbb{R}^p$ are the state and output variables of the TVF, respectively. Matrices A_i^f and C_i^f are the state matrix and output matrix of the TVF system, respectively. The dynamics of the virtual leader and the TVF determines the desired movement process of the MAS, when completing different tasks.

The n augmented systems for cooperative heterogeneous formation control are given by

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{F}_i v_i \\ \dot{v}_i &= A_i^v v_i, \\ e_i &= \tilde{C}_i x_i - C_i^v v_i, \end{aligned} \quad i = 1, 2, \dots, n, \tag{5}$$

where $v_i = \text{col}(x_0, \omega, f_i)$ is the state variable of i th reference system, and e_i is the formation output error for i th agent. The system matrices in (5) are defined as $\tilde{F}_i = [0, \tilde{D}_i, 0]$, $A_i^v = \text{blkdiag}(A_0, A_\omega, A_i^f)$, and $C_i^v = [C_0, 0, C_i^f]$.

Throughout this article, the following assumptions are needed.

Assumption 1. The eigenvalues of matrix A_i^v lie on the imaginary axis.

Assumption 2. Assume that the communication graph G of the heterogeneous multi-agent system has a spanning tree, where the virtual leader is the globally reachable root node.

Assumption 3. Assume that the pair $(\tilde{A}_i, \tilde{B}_i)$ is stabilizable, and the pairs (A_0, C_0) and $(\tilde{A}_i, \tilde{C}_i)$ are all detectable.

Assumption 4. For any $\lambda \in \sigma(A_i^v)$,

$$\text{rank} \begin{bmatrix} A_i - \lambda I_{n_i} & B_i \\ C_i & 0 \end{bmatrix} = n_i + p_i, \quad i = 1, 2, \dots, n,$$

where $\sigma(A_i^v)$ denotes the spectrum of A_i^v .

Remark 1. Note that Assumption 1 can ensure that reference systems are marginally stable. Otherwise, the reference trajectories for agents converge to the origin or increase exponentially to infinity. Assumption 1 is reasonable for MASs to track practical trajectories in a TVF without causing agent collisions under disturbances. Similar assumptions have been adopted in References 39-42. The necessary and sufficient conditions for information consensus of multi-agent systems are given in Assumption 2. Assumption 3 is made so that a dynamic state and output feedback control can be achieved. The purpose of Assumption 4 is to ensure the existence of internal model systems to achieve the output consensus for the heterogeneous MAS.¹² Meanwhile, Assumption 4 indicates that there is a unique solution (Π_i^v, U_i^v) for the following cooperative output regulation equation,

$$\begin{aligned}\Pi_i^x A_i^v &= \tilde{A}_i \Pi_i^x + \tilde{B}_i U_i^v + \tilde{F}_i, \\ \Pi_i^z A_i^v &= \Sigma_{i1} \Pi_i^z + \Sigma_{i2} (\tilde{C}_i \Pi_i^x - C_i^v), \\ 0 &= \tilde{C}_i \Pi_i^x - C_i^v,\end{aligned}\quad (6)$$

where $\Pi_i^v = \text{col}(\Pi_i^x, \Pi_i^z)$, and $(\Sigma_{i1}, \Sigma_{i2})$ is an internal model pair of A_i^v .¹²

To develop our main results in the next section, we introduce the following definitions and problems.

Problem 1. The optimal robust formation output control problem is to design a controller to minimize the objective function

$$\min J_i = \int_0^\infty (\tilde{\delta}_i^T Q_i \tilde{\delta}_i + \tilde{u}_i^T R_i \tilde{u}_i) dt, \quad i = 1, 2, \dots, n \quad (7)$$

subject to system (5) under uncertainties and disturbance for given positive definite matrices Q_i and R_i , where $\delta_i = \text{col}(x_i, z_i)$ is the augmented state variable with internal dynamics z_i , and $\tilde{\delta}_i = \delta_i - \Pi_i^v v_i$ and $\tilde{u}_i = u_i - U_i^v v_i$ are the internal model-based augmented state error and augmented input variable for optimal robust formation output control, respectively.

Problem 2. The optimal robust collision-free formation output control problem is to design a controller to minimize the objective function

$$\min J_i = \int_0^\infty (\tilde{\delta}_i^T Q_i \tilde{\delta}_i + \tilde{u}_i^T R_i \tilde{u}_i + F(\tilde{\delta}_i, \Omega_i)) dt, \quad i = 1, 2, \dots, n \quad (8)$$

subject to system (5) under uncertainties and disturbance for given positive definite matrices Q_i and R_i , where $F(\tilde{\delta}_i, \Omega_i)$ is the cost function for collision avoidance, and $\Omega_i = \{o_1, o_2, \dots, o_l\}$ is the set of l colliding objects for i th agent. The set Ω includes other agents and obstacles in the environment.

Remark 2. Note that the quadratic objective function for formation control of heterogeneous MASs is considered in Problem 1, which is extended by the cooperative optimal output regulation problem.⁴¹ Problem 2 involves nonquadratic nonlinear objective function, as the cost function $F(\tilde{\delta}_i, \Omega_i)$ for collision avoidance is usually difficult to describe in a quadratic form.

To achieve optimal robust formation control under constraints, the overview of the learning-based heterogeneous multi-agent system design is illustrated in Figure 1. An exosystem consists of a virtual leader and a disturbance system is considered. The so-called exosystem refers to an external system including reference signals to track and disturbances to reject, which is first proposed in the output regulation problem.¹² The exosystem considered here represents an external common goal to track the trajectory of the virtual leader under disturbances. Generally, the information of the exosystem is not known globally, and a distributed adaptive observer is proposed to construct an estimated exosystem based on the collaborative error obtained from communication interaction. The observer parameters are found by either an observed model-based or a model-free off-policy RL algorithm. Taking the formation deviation signal into consideration for adapting to tasks, n designed TVFs and an estimated exosystem constitute n reference systems, which pass the reference output to the optimal robust formation controller of each agent. The optimal robust formation output controller is designed with two functional modules, robust formation regulator and collision avoidance. Another model-free off-policy RL algorithm is proposed to find the optimal parameters of the controller. The unknown heterogeneous MAS realizes the optimal formation under the learning-based controller and interacts with the output information through the network.

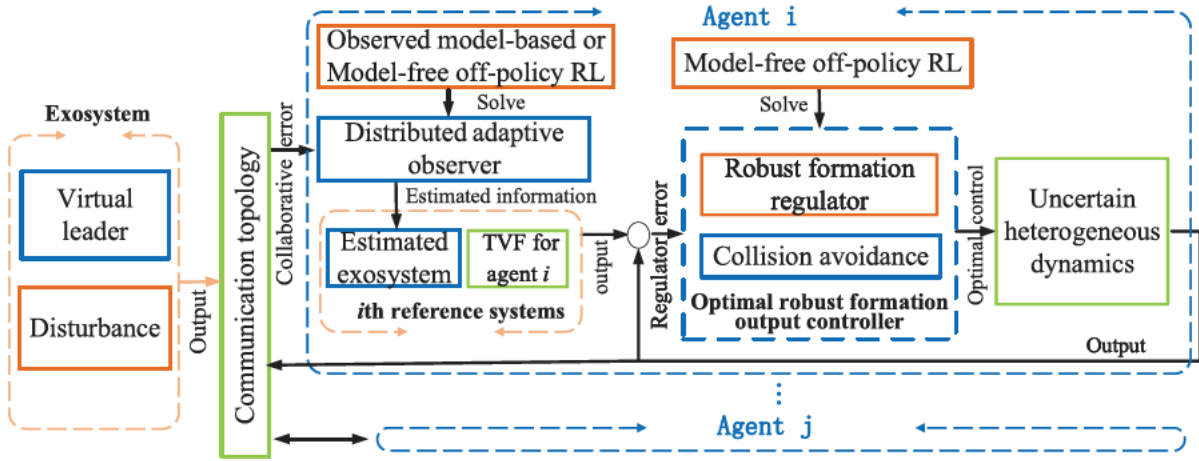


FIGURE 1 Overview of the learning-based heterogeneous multi-agent system design

3 | MAIN RESULTS

In this section, the distributed adaptive observer is proposed with two RL algorithms. Then, the two functional modules with another RL algorithm are designed for heterogeneous MASs to solve Problems 1 and 2, respectively.

3.1 | Distributed adaptive observer by RL

The exosystem consisting of the virtual leader and disturbance is recast as

$$\dot{\xi} = A^\xi \xi, \quad z^\xi = C^z \xi, \quad y^\xi = C^\xi \xi, \tag{9}$$

where $\xi = \text{col}(x_0, \omega) \in \mathbb{R}^{(n_0+n_\omega)}$, $z^\xi \in \mathbb{R}^z$, and $y^\xi \in \mathbb{R}^p$ represent the state, measurement, and output variables. Matrices $A^\xi = \text{blkdiag}(A_0, A_\omega)$, $C^z = [C_0, I]$, and $C^\xi = [C_0, 0]$ are the state matrix, measurement matrix, and output matrix of exosystem, respectively.

Based on Assumption 2, the virtual leader can be regarded as 0th agent, and the whole Laplacian matrix becomes

$$L_{n+1} = \begin{bmatrix} 0 & 0_{1 \times n} \\ L_0 & L_1 \end{bmatrix}, \tag{10}$$

where $L_1 = L + A_{n0}$, and L is the Laplacian matrix of n heterogeneous MASs. The adjacency matrix between the virtual leader and the n agents is $A_{n0} = \text{diag}[a_{i0}]$.

We design the distributed adaptive observer as

$$\begin{aligned} \dot{\hat{\xi}}_i &= \hat{A}_i^\xi \hat{\xi}_i - K_{oi} \theta_i \psi_i, \quad K_{oi} = P_{oi} (\hat{C}_i^z)^T, \\ \dot{\hat{A}}_i^\xi &= -c \left(\sum_{j=1}^n a_{ij} (\hat{A}_i^\xi - \hat{A}_j^\xi) + a_{i0} (\hat{A}_i^\xi - A^\xi) \right), \quad c > 0, \\ \dot{\hat{C}}_i^z &= -c \left(\sum_{j=1}^n a_{ij} (\hat{C}_i^z - \hat{C}_j^z) + a_{i0} (\hat{C}_i^z - C^z) \right), \quad c > 0, \\ \dot{\hat{C}}_i^\xi &= -c \left(\sum_{j=1}^n a_{ij} (\hat{C}_i^\xi - \hat{C}_j^\xi) + a_{i0} (\hat{C}_i^\xi - C^\xi) \right), \quad c > 0, \\ \dot{X}_i &= (\hat{A}_i^\xi)^T X_i + (\hat{C}_i^z)^T U_i = \left((\hat{A}_i^\xi)^T - (\hat{C}_i^z)^T \hat{C}_i^\xi P_{oi} \right) X_i, \end{aligned}$$

$$\begin{aligned}\psi_i &= \sum_{j=1}^n a_{ij} \left(\hat{C}_i^z \hat{\xi}_i - \hat{C}_j^z \hat{\xi}_j \right) + a_{i0} \left(\hat{C}_i^z \hat{\xi}_i - z^{\xi} \right), \\ \theta_i &= d\psi_i^T \psi_i, \quad d > 0, \theta_i(0) > 1,\end{aligned}\quad (11)$$

where matrices \hat{A}_i^{ξ} , \hat{C}_i^z , and \hat{C}_i^{ξ} are the estimations of state matrix, measurement matrix, and output matrix of the exosystem. Variable $\hat{\xi}_i$ is the observation of ξ , and X_i is a virtual auxiliary variable. The symmetric positive definite matrix P_{oi} is gained by Algorithm 1 or Algorithm 2 under given symmetric positive definite matrix Q .

Algorithm 1. The observed model-based off-policy RL algorithm to find matrix P_{oi}

Step 1. Initialization (number of iterations $k = 0$): given an admissible policy $U_i^0 = -\hat{C}_i^z P_{oi}^0 X_i$

Step 2. Evaluate policy by finding P_{oi}^k using the following Bellman equation:

$$\begin{aligned}X_i^T(t + \Delta T) P_{oi}^k X_i(t + \Delta T) - X_i^T(t) P_{oi}^k X_i(t) \\ = \int_t^{t+\Delta T} - (X_i^T(\tau) Q X_i(\tau) + (U_i^k)^T U_i^k) d\tau.\end{aligned}\quad (12)$$

Update policy by

$$U_i^{k+1} = -\hat{C}_i^z P_{oi}^k X_i(t).\quad (13)$$

Step 3. Check end condition:

If $\|P_{oi}^{k+1} - P_{oi}^k\| < \eta$, where η is a small constant

Set $P_{oi} = P_{oi}^k$; Stop iteration.

else set $k = k + 1$ and go to Step 2.

Defining the observer state error variable as $\tilde{\xi}_i = \hat{\xi}_i - \xi$, and dynamical error of system matrices as $\tilde{A}_i^{\xi} = \hat{A}_i^{\xi} - A^{\xi}$, $\tilde{C}_i^z = \hat{C}_i^z - C^z$, and $\tilde{C}_i^{\xi} = \hat{C}_i^{\xi} - C^{\xi}$. We denote related column vectors and matrices of n agents by

$$\begin{aligned}\underline{\xi} &= I_n \otimes \xi, \quad \tilde{\xi} = \text{col}(\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_n), \quad \hat{\xi} = \text{col}(\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_n), \\ \underline{A}^{\xi} &= \text{diag}(A^{\xi}, A^{\xi}, \dots, A^{\xi}), \quad \tilde{A}^{\xi} = \text{diag}(\tilde{A}_1^{\xi}, \tilde{A}_2^{\xi}, \dots, \tilde{A}_n^{\xi}), \quad \hat{A}^{\xi} = \text{diag}(\hat{A}_1^{\xi}, \hat{A}_2^{\xi}, \dots, \hat{A}_n^{\xi}), \\ \underline{C}^z &= \text{diag}(C^z, C^z, \dots, C^z), \quad \tilde{C}^z = \text{diag}(\tilde{C}_1^z, \tilde{C}_2^z, \dots, \tilde{C}_n^z), \quad \hat{C}^z = \text{diag}(\hat{C}_1^z, \hat{C}_2^z, \dots, \hat{C}_n^z), \\ \underline{C}^{\xi} &= \text{diag}(C^{\xi}, C^{\xi}, \dots, C^{\xi}), \quad \tilde{C}^{\xi} = \text{diag}(\tilde{C}_1^{\xi}, \tilde{C}_2^{\xi}, \dots, \tilde{C}_n^{\xi}), \quad \hat{C}^{\xi} = \text{diag}(\hat{C}_1^{\xi}, \hat{C}_2^{\xi}, \dots, \hat{C}_n^{\xi}), \\ \Psi &= \text{diag}(\psi_1, \psi_2, \dots, \psi_n), \quad \Theta = \text{diag}(\theta_1, \theta_1, \dots, \theta_n).\end{aligned}$$

Recall the following results in order to derive our main results.

Lemma 1 (43). *If the communication topology G contains a spanning tree and the virtual leader is the root node, the matrix L_1 in (10) is a nonsingular M -matrix with positive eigenvalues.*

Lemma 2 (12). *The pair of (Σ_1, Σ_2) is the p -copy internal model of square matrix A , if the pair has the form:*

$$\Sigma_1 = T \begin{bmatrix} S_1 & S_2 \\ 0 & \bar{\Sigma}_1 \end{bmatrix} T^{-1}, \quad \Sigma_2 = T \begin{bmatrix} S_3 \\ \bar{\Sigma}_2 \end{bmatrix},\quad (14)$$

where T is any nonsingular matrix, S_i , $i = 1, 2, 3$ are any constant matrices, and for any $p > 0$,

$$\bar{\Sigma}_1 = \text{blkdiag}(\alpha_{11}, \dots, \alpha_{1p}), \quad \bar{\Sigma}_2 = \text{blkdiag}(\alpha_{21}, \dots, \alpha_{2p}).$$

If the minimal polynomial of A is

$$\lambda^n + a_1\lambda^{(n-1)} + \dots + a_{(n-1)}\lambda + a_n$$

then $(\alpha_{1j}, \alpha_{2j})$ is controllable in the form of

$$\alpha_{1j} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \\ -a_n & -a_{(n-1)} & \dots & -a_1 \end{bmatrix}, \quad \alpha_{2j} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

where $j = 1, 2, \dots, p$.

The first main result of this article is given as follows

Theorem 1. Consider systems (1)–(4) satisfying Assumptions 1–4. The observer dynamical error and state error for exosystem (9) converge to zero under observers (11) solved by Algorithm 1 or Algorithm 2.

Proof. We first prove the convergence of the dynamical error. Based on observer (11), differentiating the dynamical error shows

$$\begin{aligned} \text{vec}(\dot{\tilde{A}}^\xi) &= \text{vec}(\dot{\hat{A}}^\xi) - \text{vec}(\underline{\hat{A}}^\xi) = -cL_1 \otimes \text{Ivec}(\hat{A}^\xi), \\ \text{vec}(\dot{\tilde{C}}^z) &= \text{vec}(\dot{\hat{C}}^z) - \text{vec}(\underline{\hat{C}}^z) = -cL_1 \otimes \text{Ivec}(\hat{C}^z), \\ \text{vec}(\dot{\tilde{C}}^\xi) &= \text{vec}(\dot{\hat{C}}^\xi) - \text{vec}(\underline{\hat{C}}^\xi) = -cL_1 \otimes \text{Ivec}(\hat{C}^\xi), \end{aligned} \tag{15}$$

where $\text{vec}()$ represents the matrix to be reshaped into a variable by column.

From Lemma 1, the eigenvalues of L_1 have positive real parts. Since $c > 0$, all the eigenvalues of matrix $-cL_1 \otimes I$ locate in the left half of the S plane. Therefore the dynamical error matrix after reshaping converges to zero as

$$\begin{aligned} \lim_{t \rightarrow \infty} \text{vec}(\tilde{A}^\xi) &= 0, & \lim_{t \rightarrow \infty} \text{vec}(\tilde{A}_i^\xi) &= 0, & \lim_{t \rightarrow \infty} \tilde{A}_i^\xi &= 0, \\ \lim_{t \rightarrow \infty} \text{vec}(\tilde{C}^z) &= 0, & \lim_{t \rightarrow \infty} \text{vec}(\tilde{C}_i^z) &= 0, & \lim_{t \rightarrow \infty} \tilde{C}_i^z &= 0, \\ \lim_{t \rightarrow \infty} \text{vec}(\tilde{C}^\xi) &= 0, & \lim_{t \rightarrow \infty} \text{vec}(\tilde{C}_i^\xi) &= 0, & \lim_{t \rightarrow \infty} \tilde{C}_i^\xi &= 0. \end{aligned} \tag{16}$$

The dynamical errors converge to zero exponentially at least at the rate of $c\lambda_{\min}(L_1)$, where $\lambda_{\min}(L_1)$ represents the minimum eigenvalues of matrix L_1 . The rate of convergence can be accelerated by choosing a sufficiently large c . Then, for the convergence of the observation state error from (11), it follows that

$$\begin{aligned} \dot{\xi}_i &= \hat{A}_i^\xi \hat{\xi}_i - A^\xi \xi + \theta_i K_{oi} \sum_{j=0}^N a_{ij} \left(\hat{C}_j^z \hat{\xi}_j - \hat{C}_i^z \hat{\xi}_i \right) \\ &= A^\xi \hat{\xi}_i - A^\xi \xi + \hat{A}_i^\xi \hat{\xi}_i - A^\xi \hat{\xi}_i + \theta_i K_{oi} \sum_{j=0}^N a_{ij} \left((\tilde{C}_j^z + C^z) \hat{\xi}_j - (\tilde{C}_i^z + C^z) \hat{\xi}_i \right) \\ &= A^\xi \tilde{\xi}_i + \hat{A}_i^\xi \hat{\xi}_i + \theta_i K_{oi} C^z \sum_{j=0}^N a_{ij} (\tilde{\xi}_j - \tilde{\xi}_i) + \theta_i K_{oi} \sum_{j=0}^N a_{ij} (\tilde{C}_j^z (\tilde{\xi}_j + \xi) - \tilde{C}_i^z (\tilde{\xi}_i + \xi)) \\ &= A^\xi \tilde{\xi}_i + \hat{A}_i^\xi \tilde{\xi}_i + \hat{A}_i^\xi \xi + \theta_i P_0 (C^z)^T C^z \sum_{j=0}^N a_{ij} (\tilde{\xi}_j - \tilde{\xi}_i) + \theta_i \tilde{K}_{oi} C^z \sum_{j=0}^N a_{ij} (\tilde{\xi}_j - \tilde{\xi}_i) \\ &\quad + \theta_i K_{oi} \sum_{j=0}^N a_{ij} (\tilde{C}_j^z \tilde{\xi}_j - \tilde{C}_i^z \tilde{\xi}_i) + \theta_i K_{oi} \sum_{j=0}^N a_{ij} (\tilde{C}_j^z - \tilde{C}_i^z) \xi \end{aligned}$$

$$\begin{aligned}
&= A^\xi \tilde{\xi}_i + \theta_i P_0 (C^\zeta)^T C^\zeta \sum_{j=0}^N a_{ij} (\tilde{\xi}_j - \tilde{\xi}_i) + \tilde{A}_i^\xi \xi + \theta_i K_{oi} \sum_{j=0}^N a_{ij} (\tilde{C}_j^\zeta - \tilde{C}_i^\zeta) \xi \\
&\quad + \tilde{A}_i^\xi \tilde{\xi}_i + \theta_i \tilde{K}_{oi} C^\zeta \sum_{j=0}^N a_{ij} (\tilde{\xi}_j - \tilde{\xi}_i) + \theta_i K_{oi} \sum_{j=0}^N a_{ij} (\tilde{C}_j^\zeta \tilde{\xi}_j - \tilde{C}_i^\zeta \tilde{\xi}_i), \tag{17}
\end{aligned}$$

where $\hat{C}_0^\zeta = C^\zeta$, $\hat{\xi}_0 = \xi$, $\tilde{K}_{oi} = P_{oi}(\hat{C}_i^\zeta)^T - P_0(C^\zeta)^T$, and P_0 is the solution of the following Riccati equation

$$P_0 A^\xi + (A^\xi)^T P_0 - P_0 (C^\zeta)^T C^\zeta P_0 + Q = 0. \tag{18}$$

Note that the pair (A^ξ, C^ζ) in (9) is stabilizable under Assumption 3. Therefore, the Riccati equation (18) admits a unique positive definite matrix P_0^* by given a positive definite matrix $Q = Q^{-1} > 0$. However, both A^ξ and C^ζ are not known for each agent as global model information. Although the dynamical model is estimated by (11), it is difficult to guarantee that the observation pair $(\hat{A}_i^\xi, \hat{C}_i^\zeta)$ is stabilizable at every moment.

In order to overcome the difficulties, a virtual dual system is introduced as

$$\dot{X}_i = (\hat{A}_i^\xi)^T X_i + (\hat{C}_i^\zeta)^T U_i = \left((\hat{A}_i^\xi)^T - (\hat{C}_i^\zeta)^T \hat{C}_i^\zeta P_{oi} \right) X_i. \tag{19}$$

The Bellman equation for system (19) is

$$\begin{aligned}
&X_i^T(t + \Delta T) P_{oi}^k X_i(t + \Delta T) - X_i^T(t) P_{oi}^k X_i(t) \\
&= \int_t^{t+\Delta T} X_i^T(t) \left[\left((\hat{A}_i^\xi)^T - (\hat{C}_i^\zeta)^T \hat{C}_i^\zeta P_{oi}^k \right)^T P_{oi}^k + P_{oi}^k \left((\hat{A}_i^\xi)^T - (\hat{C}_i^\zeta)^T \hat{C}_i^\zeta P_{oi}^k \right) \right] X_i(t) d\tau \\
&= \int_t^{t+\Delta T} - (X_i^T(t) Q X_i(t) + (U_i^k)^T U_i^k) d\tau. \tag{20}
\end{aligned}$$

Based on Theorem 4⁴⁴ and Lemma 4,⁴⁵ the matrix P_{oi} from Algorithm 1 converges to the solution P_{oi}^* of the following Riccati equation

$$P_{oi} \hat{A}_i^\xi + (\hat{A}_i^\xi)^T P_{oi} - P_{oi} (\hat{C}_i^\zeta)^T \hat{C}_i^\zeta P_{oi} + Q = 0. \tag{21}$$

To remove the model information \hat{C}_i^ζ in Algorithm 1, the Bellman equation (20) can be rewritten as

$$\begin{aligned}
&X_i^T(t + \Delta T) P_{oi}^k X_i(t + \Delta T) - X_i^T(t) P_{oi}^k X_i(t) \\
&= \int_t^{t+\Delta T} \left[-X_i^T(t) \left(Q_i + (K_{oi}^k)^T K_{oi}^k \right) X_i(t) + 2(U_i + K_{oi}^k X_i(t))^T \hat{C}_i^\zeta P_{oi}^k X_i(t) \right] d\tau \\
&= \int_t^{t+\Delta T} \left[-X_i^T(t) \left(Q_i + (K_{oi}^k)^T K_{oi}^k \right) X_i(t) + 2(U_i - U_i^k)^T K_{oi}^{k+1} X_i(t) \right] d\tau. \tag{22}
\end{aligned}$$

From (22), the model-free algorithm is given by Algorithm 2 and the obtained P_{oi} converges to the solution P_{oi}^* of the Riccati equation (21).^{42,45}

If $\lim_{t \rightarrow \infty} \hat{A}_i^\xi = 0$, $\lim_{t \rightarrow \infty} \tilde{C}_i^\zeta = 0$, and P_0^* is the unique, it can be extracted by Lemma 2.2 and Theorem 3.1 in Reference 22 that $\lim_{t \rightarrow \infty} \tilde{P}_{oi} = \lim_{t \rightarrow \infty} (P_{oi} - P_0^*) = 0$. It implies that $\lim_{t \rightarrow \infty} \tilde{K}_{oi} = 0$.

The derivatives of state error $\tilde{\xi}$ is

$$\begin{aligned}
\dot{\tilde{\xi}} &= (I_n \otimes A^\xi - \Theta (L_1 \otimes (P_0(C^\zeta)^T C^\zeta))) \tilde{\xi} \\
&\quad + \left(\hat{A}^\xi - \Theta \tilde{K}_o (L_1 \otimes C^\zeta) - \Theta K_o (L_1 \otimes I_z) \tilde{C}^\zeta \right) \tilde{\xi} \\
&\quad + \left(\hat{A}^\xi - \Theta K_o (L_1 \otimes I_z) \tilde{C}^\zeta \right) \underline{\xi} \\
&= F_0 \tilde{\xi} + F_1 \underline{\xi} + F_2, \tag{23}
\end{aligned}$$

where

$$\begin{aligned}
\bar{K}_o &= \text{diag}(\bar{K}_{o1}, \bar{K}_{o2}, \dots, \bar{K}_{on}) \quad K_o = \text{diag}(K_{o1}, K_{o2}, \dots, K_{on}), \\
F_0 &= I_n \otimes A^\xi - \Theta(L_1 \otimes (P_0(C^\xi)^T C^\xi)), \\
F_1 &= \bar{A}^\xi - \Theta \bar{K}_o(L_1 \otimes C^\xi) - \Theta K_o(L_1 \otimes I_z) \tilde{C}^\xi, \\
F_2 &= \left(\bar{A}^\xi - \Theta K_o(L_1 \otimes I_z) \tilde{C}^\xi \right) \underline{\xi}.
\end{aligned} \tag{24}$$

Under Assumptions 1–3 and observer designed in (11), system $\dot{\tilde{\xi}} = F_0 \tilde{\xi}$ is exponentially stable based on Lemma 2.1.²² Therefore, system (23) is input-to-state stable with $F_1 \tilde{\xi} + F_2$ as the input. Since $\lim_{t \rightarrow \infty} \bar{A}^\xi = 0$, $\lim_{t \rightarrow \infty} \tilde{C}^\xi = 0$ and $\lim_{t \rightarrow \infty} \bar{K}_o = 0$ hold, functions F_1 and F_2 are bounded and decay to zero exponentially under Assumptions 1 and 2. Based on Lemma 1,⁴⁶ if F_1 and F_2 tend to zero exponentially, so does $\tilde{\xi}$. It turns out that

$$\lim_{t \rightarrow \infty} \tilde{\xi}_i = 0.$$

This completes the proof. ■

Algorithm 2. The model-free off-policy RL algorithm to find matrix P_{oi}

Step 1. Data collection: For agent i , apply a fixed admissible control policy $U_i^k = -K_{oi}^0 X_i + e_{di}$, and collect the system information such as input and output. e_{di} is a noise term.

Step 2. Date reuse for solution: Set $k = 0$

Step 2.1. Evaluate policy by solve P_{oi}^k, K_{oi}^{k+1} of the following Bellman equation using the information from Step 1.

$$\begin{aligned}
& X_i^T(t + \Delta T) P_{oi}^k X_i(t + \Delta T) - X_i^T(t) P_{oi}^k X_i(t) \\
&= \int_t^{t+\Delta T} \left[-X_i^T(t) \left(Q_i + (K_{oi}^k)^T K_{oi}^k \right) X_i(t) + 2(U_i - U_i^k)^T (K_{oi}^{k+1} X_i(t)) \right] d\tau.
\end{aligned} \tag{25}$$

Step 2.2. Update control policy by

$$U_i^{k+1} = -K_{oi}^{k+1} X_i(t). \tag{26}$$

Step 2.3. Check end condition:

If $\|P_{oi}^{k+1} - P_{oi}^k\| < \eta$, where η is a small constant

Set $P_{oi} = P_{oi}^k$ and $K_{oi} = K_{oi}^k$; Stop iteration.

else set $k = k + 1$ and go to Step 2.1.

Remark 3. Note that the admissible policy can be achieved by choosing nonzero initial values of \hat{C}_i^ξ and P_{oi}^0 , and the Bellman equation (12) can be solved by the least-squares method under some persistence of excitation (PE) condition.^{42,47} Algorithm 1 is based on estimated model information \hat{C}_i^ξ , while Algorithm 2 is model-free. Note that the Assumption 1 can be relaxed if c and initial P_{oi}^0 are selected large enough.^{44,46} The adaptive parameter θ is to ensure system stability after relaxing Assumption 1.⁴⁸

Remark 4. Although a great amount of work has been devoted to the adaptive observer design to estimate the exosystem, prior global knowledge on communication^{19,20} or partial model information^{13,21} is usually involved in the observer design, and it is difficult to guarantee the solvability of the Riccati equation (21) with estimated models at each time. Both global topology information and model information is removed by the developed adaptive observer, and the Riccati equation (21) can be solved by the RL-based iterative technology.

3.2 | Optimal robust formation output control

For the formation control considering the TVF dynamics in (4), the n estimated reference systems for n agents are

$$\hat{v}_i = \hat{A}_i^v \hat{v}_i, \quad \hat{y}_i^v = \hat{C}_i^v \hat{v}_i, \quad i = 1, 2, \dots, n, \tag{27}$$

where $\hat{v}_i = \text{col}(\hat{\xi}_i, f_i)$, $\hat{A}_i^v = \text{blkdiag}(\hat{A}_i^{\xi}, A_i^f)$, and $\hat{C}_i^v = [\hat{C}_i^{\xi}, C_i^f]$.

We design the state feedback control law as

$$\begin{aligned} u_i &= -K_i^x x_i - K_i^z z_i, \\ z_i &= \hat{\Sigma}_{i1} z_i + \Sigma_{i2} \hat{e}_i, \\ \hat{e}_i &= y_i - \hat{C}_i^{\xi} \hat{\xi}_i - C_i^f f_i, \end{aligned} \quad (28)$$

where z_i is the state of the dynamic compensator. $(\hat{\Sigma}_{i1}, \Sigma_{i2})$ is the p -copy internal model pair of \hat{A}_i^v . From Theorem 1, it turns out that

$$\begin{aligned} \lim_{t \rightarrow \infty} \bar{A}_i^v &= \lim_{t \rightarrow \infty} (\hat{A}_i^v - A_i^v) = 0, & \lim_{t \rightarrow \infty} \tilde{C}_i^v &= \lim_{t \rightarrow \infty} (\hat{C}_i^v - C_i^v) = 0, \\ \lim_{t \rightarrow \infty} \tilde{\Sigma}_{i1} &= \lim_{t \rightarrow \infty} (\hat{\Sigma}_{i1} - \Sigma_{i1}) = 0, & \lim_{t \rightarrow \infty} \tilde{v}_i &= \lim_{t \rightarrow \infty} (\hat{v}_i - v_i) = 0, \end{aligned} \quad (29)$$

where $K_i = [K_i^x, K_i^z]$ is the optimal control law from Algorithm 3, and $(\Sigma_{i1}, \Sigma_{i2})$ is p -copy internal model pair of $\text{blkdiag}(A_i^{\xi}, A_i^f)$. Note that Σ_{i2} is constructed without observation error due to its definition in (14).

The second main result of this article is given as follows.

Algorithm 3. The model-free off-policy RL algorithm for optimization

Step 1. Data collection: For agent i , apply a fixed admissible control policy $u_i^k = -K_i^0 \delta_i + e_{di}$, and collect the system information such as input and output. e_{di} is a noise term.

Step 2. Date reuse for solution: Set $k = 0$

Step 2.1. Evaluate policy by solve P_i^k , K_i^{k+1} , and $P_i^k \tilde{B}_i^c$ of the following Bellman equation using the information from Step 1.

$$\begin{aligned} & \delta_i^T(t + \Delta T) P_i^k \delta_i(t + \Delta T) - \delta_i^T(t) P_i^k \delta_i(t) \\ &= \int_t^{t+\Delta T} [-\delta_i^T(\tau) (Q_i + (K_i^k)^T R_i K_i^k) \delta_i(\tau) + 2(u_i - u_i^k)^T R_i K_i^{k+1} \delta_i(\tau) + 2\tilde{v}_i^T (P_i^k \tilde{B}_i^c)^T \delta_i(\tau)] d\tau. \end{aligned} \quad (30)$$

Step 2.2. Update control policy by

$$u_i^{k+1} = -K_i^{k+1} \delta_i(t). \quad (31)$$

Step 2.3. Check end condition:

If $\|K_i^k - K_i^{k-1}\| < \eta$, where η is a small constant

Set $P_i = P_i^k$ and $K_i = K_i^k$; Stop iteration.

Else

set $k = k + 1$ and go to Step 2.1.

Theorem 2. Consider systems (1)–(4) satisfying Assumptions 1–4. The optimal robust formation output control problem defined in Problem 1 is solved by the adaptive observers in (11) using Algorithm 2 and the optimal controller in (28) by Algorithm 3.

Proof. Substitution of feedback law (28) into the augmented system (5) yields

$$\begin{aligned} \begin{bmatrix} \dot{x}_i \\ \dot{z}_i \end{bmatrix} &= \begin{bmatrix} \hat{A}_i & 0 \\ \Sigma_{i2} \tilde{C}_i & \hat{\Sigma}_{i1} \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} + \begin{bmatrix} \tilde{B}_i \\ 0 \end{bmatrix} u_i + \begin{bmatrix} \tilde{F}_i v_i \\ -\Sigma_{i2} \hat{C}_i \hat{v}_i \end{bmatrix} \\ &= \begin{bmatrix} \hat{A}_i & 0 \\ \Sigma_{i2} \tilde{C}_i & \Sigma_{i1} \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} \tilde{B}_i \\ 0 \end{bmatrix} \begin{bmatrix} K_i^x & K_i^z \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} + \begin{bmatrix} \tilde{F}_i \\ -\Sigma_{i2} C_i^v \end{bmatrix} \hat{v}_i + \begin{bmatrix} -\tilde{F}_i \tilde{v}_i \\ \hat{\Sigma}_{i1} z_i - \Sigma_{i2} (\tilde{C}_i^v v_i + \tilde{C}_i^v \tilde{v}_i) \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
 &= \begin{bmatrix} \tilde{A}_i - \tilde{B}_i K_i^x & -\tilde{B}_i K_i^z \\ \Sigma_{i2} \tilde{C}_i & \Sigma_{i1} \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} + \begin{bmatrix} \tilde{F}_i \\ -\Sigma_{i2} C_i^v \end{bmatrix} v_i + \begin{bmatrix} 0 \\ \tilde{\Sigma}_{i1} z_i - \Sigma_{i2} (C_i^v \tilde{v}_i + \tilde{C}_i^v v_i + \tilde{C}_i^v \tilde{v}_i) \end{bmatrix}, \\
 e_i &= \begin{bmatrix} \tilde{C}_i & 0 \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} - C_i^v v_i = \left(\begin{bmatrix} \tilde{C}_i & 0 \end{bmatrix} \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \hat{C}_i^v \hat{v}_i \right) - (-\hat{C}_i^v \tilde{v}_i - \tilde{C}_i^v v_i) = \hat{e}_i - \tilde{e}_i.
 \end{aligned} \tag{32}$$

Define augmented state variable as $\delta_i = \text{col}(x_i, z_i)$, system (32) is rewritten by

$$\begin{aligned}
 \delta_i &= \tilde{A}_i^\delta \delta_i + \tilde{B}_i^\delta u_i + \tilde{B}_i^c \hat{v}_i + M_{i1} \\
 &= \tilde{A}_i^c \delta_i + \tilde{B}_i^c v_i + M_{i2}, \\
 e_i &= \tilde{C}_i^c \delta_i - C_i^v v_i, \\
 \tilde{e}_i &= -\hat{C}_i^v \tilde{v}_i - \tilde{C}_i^v v_i,
 \end{aligned} \tag{33}$$

where

$$\begin{aligned}
 \tilde{A}_i^\delta &= \begin{bmatrix} \tilde{A}_i & 0 \\ \Sigma_{i2} \tilde{C}_i & \Sigma_{i1} \end{bmatrix}, \quad \tilde{B}_i^\delta = \begin{bmatrix} \tilde{B}_i \\ 0 \end{bmatrix}, \quad \tilde{B}_i^c = \begin{bmatrix} \tilde{F}_i \\ -\Sigma_{i2} C_i^v \end{bmatrix}, \quad M_{i1} = \begin{bmatrix} -\tilde{F}_i \tilde{v}_i \\ \tilde{\Sigma}_{i1} z_i - \Sigma_{i2} (\tilde{C}_i^v v_i + \tilde{C}_i^v \tilde{v}_i) \end{bmatrix}, \\
 \tilde{A}_i^c &= \begin{bmatrix} \tilde{A}_i - \tilde{B}_i K_i^x & -\tilde{B}_i K_i^z \\ \Sigma_{i2} \tilde{C}_i & \Sigma_{i1} \end{bmatrix}, \quad M_{i2} = \begin{bmatrix} 0 \\ \tilde{\Sigma}_{i1} z_i - \Sigma_{i2} (C_i^v \tilde{v}_i + \tilde{C}_i^v v_i + \tilde{C}_i^v \tilde{v}_i) \end{bmatrix}, \quad \tilde{C}_i^c = [\tilde{C}_i \ 0].
 \end{aligned}$$

According to Lemmas 1.20 and 1.27 in Reference 12, if A_i^c without uncertainties is Hurwitz, then \tilde{A}_i^c is also Hurwitz under Assumption 4, for any $\Delta \star_i \in W$. If $(\Sigma_{i1}, \Sigma_{i2})$ is the p-copy internal model pair of A_i^v , for any $\Delta \star_i \in W$, the following equations

$$\Pi_i^x A_i^v = (\tilde{A}_i - \tilde{B}_i K_i^x) \Pi_i^x - \tilde{B}_i K_i^z \Pi_i^z + \tilde{F}_i, \quad \Pi_i^z A_i^v = \Sigma_{i1} \Pi_i^x + \Sigma_{i2} (\tilde{C}_i \Pi_i^x - C_i^v) \tag{34}$$

have a unique solution (Π_i^x, Π_i^z) , which satisfies $0 = \tilde{C}_i \Pi_i^x - C_i^v$.

Constructing a variable $\Pi_i^v = \text{col}(\Pi_i^x, \Pi_i^z)$, and substituting it into (34) leads to

$$\Pi_i^v A_i^v = \tilde{A}_i^c \Pi_i^v + \tilde{B}_i^c, \quad 0 = \tilde{C}_i^c \Pi_i^v - C_i^v. \tag{35}$$

Equation (35) is equivalent to cooperative output regulation (6), when $U_i^v = -K_i^x \Pi_i^x - K_i^z \Pi_i^z$. From system (5) and (35), the dynamics of the state error $\tilde{\delta}_i = \delta_i - \Pi_i^v v_i$ is

$$\dot{\tilde{\delta}}_i = \tilde{A}_i^c \tilde{\delta}_i + \tilde{B}_i^c v_i + M_{i2} - \Pi_i^v A_i^v v_i = \tilde{A}_i^c \tilde{\delta}_i + M_{i2}. \tag{36}$$

Similarly, the formation output error is rewritten as

$$e_i = \tilde{C}_i^c \tilde{\delta}_i - C_i^v v_i = \tilde{C}_i^c \tilde{\delta}_i + (\tilde{C}_i^c \Pi_i^v - C_i^v) v_i = \tilde{C}_i^c \tilde{\delta}_i. \tag{37}$$

The closed-loop formation error of the heterogeneous MAS can be expressed as

$$\dot{\tilde{\delta}}_i = \tilde{A}_i^c \tilde{\delta}_i + M_{i2}, \quad e_i = \tilde{C}_i^c \tilde{\delta}_i.$$

From Theorem 1, it indicates that M_{i1} , M_{i2} , and \tilde{e}_i converges to zero as t tends to infinity. The convergence rate is accelerated by increasing c . It is easy to see that $\lim_{t \rightarrow \infty} \tilde{\delta}_i = 0$, $\lim_{t \rightarrow \infty} e_i = 0$, and $\lim_{t \rightarrow \infty} \hat{e}_i = 0$, if \tilde{A}_i^c is Hurwitz.

Therefore, the optimization problem defined in Problem 1 can be regarded as equivalent to the optimization problem of HJB equation⁴² for the following error system

$$\dot{\tilde{\delta}}_i = \tilde{A}_i^\delta \tilde{\delta}_i + \tilde{B}_i^\delta \tilde{u}_i, \quad \tilde{u}_i = -K_i \tilde{\delta}_i, \quad e_i = \tilde{C}_i^c \tilde{\delta}_i.$$

Note

$$u_i = \tilde{u}_i + U_i^v v_i = -K_i^x \tilde{x}_i - K_i^z \tilde{z}_i + U_i^v v_i = -K_i^x x_i - K_i^z z_i = -K_i \delta_i \quad (38)$$

is consistent with the controller in (28). The optimal control law K_i is the unique solution of the following generalized algebraic Riccati equation

$$(\tilde{A}_i^\delta)^T P_i + P_i \tilde{A}_i^\delta - P_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T P_i + Q_i = 0, \quad K_i = R_i^{-1} (\tilde{B}_i^\delta)^T P_i, \quad (39)$$

where R_i and Q_i are given symmetric positive definite matrices. As the model information is unknown for agents, model-free policy iteration is obtained based on the following Bellman equation from system (33).

$$\begin{aligned} & \delta_i^T(t + \Delta T) P_i^k \delta_i(t + \Delta T) - \delta_i^T(t) P_i^k \delta_i(t) \\ &= \int_t^{t+\Delta T} \left[\delta_i^T(\tau) \left((\tilde{A}_i^\delta - \tilde{B}_i^\delta K_i^k)^T P_i + P_i (\tilde{A}_i^\delta - \tilde{B}_i^\delta K_i^k) \right) \delta_i(\tau) + 2(u_i - u_i^k)^T R_i K_i^{k+1} \delta_i(\tau) + 2\tilde{v}_i^T (\tilde{B}_i^c)^T P_i^k \delta_i \right] d\tau \\ &= \int_t^{t+\Delta T} \left[-\delta_i^T(\tau) (Q_i + (K_i^k)^T R_i K_i^k) \delta_i(\tau) + 2(u_i - u_i^k)^T R_i K_i^{k+1} \delta_i(\tau) + 2\tilde{v}_i^T (\tilde{B}_i^c)^T P_i^k \delta_i \right] d\tau \\ &= \int_t^{t+\Delta T} \left[-(\delta_i^T \otimes \delta_i^T) \text{vec} (Q_i + (K_i^k)^T R_i K_i^k) + 2 \left((\delta_i^T \otimes \delta_i^T) (I \otimes (K_i^k)^T R_i) + (\delta_i^T \otimes u_i^T) (I \otimes R_i) \right) \text{vec} (K_i^{k+1}) \right. \\ & \quad \left. + 2 (\delta_i^T \otimes \tilde{v}_i^T) \text{vec} ((\tilde{B}_i^c)^T P_i^k) \right] d\tau, \end{aligned} \quad (40)$$

where $M_{i1} = 0$ by selecting a sufficient large c .⁴¹ Expanded by Theorem 3⁴² and Theorem 2,⁴¹ the matrices P_i and K_i from Algorithm 3 converge to the optimal solutions P_i^* and K_i^* of the Riccati equation (39). The unique solution for K_i^{k+1} , P_i^k , and $(\tilde{B}_i^c)^T P_i^k$ from the Bellman equation (40) can be solved by the least squares method under some PE condition for full column rank, which can be achieved with enough samples in Step 1.

It can be seen from (39) that the solution of Algorithm 3 not only ensures that matrices \tilde{A}_i^c is Hurwitz, but also minimizes the objective function in (7). Furthermore, as $\lim_{t \rightarrow \infty} \tilde{\delta}_i = 0$ and $\lim_{t \rightarrow \infty} \tilde{e}_i = 0$, the measured formation output error $\lim_{t \rightarrow \infty} \hat{e}_i = 0$. Without knowing the uncertain system matrices \tilde{A}_i^δ and \tilde{B}_i^δ , Algorithms 2 and 3 solve Problem 1, which completes the proof. ■

Remark 5. Note that the p -copy internal model-based variable z_i in (28) is designed as a dynamic compensator to compensate the influence of system uncertainties, assist the heterogeneous MAS to follow the virtual leader, and reject the disturbance, where p is the dimension of outputs of systems (1) and (27). Consequently, the heterogeneous MAS achieves the optimal formation even under system uncertainties and disturbances.

Remark 6. It is worth noting that all Algorithms 1–3 are based on off-policy RL, in which the target policy differs from the behavior policy. Observer-model based Algorithm 1 and model-free Algorithm 2 aim to achieve the stability of adaptive observer (11), while model-free Algorithm 3 is designed to solve the formation optimization Problem 1. Compared with Bellman equations and steps in the three algorithms, a noise term e_{di} is introduced in Algorithms 2 and 3 for stimulating deviation of the Bellman equations (25) and (30) without requiring any model information. In addition to matrix P_{oi} in Algorithm 1, both matrices P_{oi} and K_{oi} are collected in Algorithm 2, and parameters $(P_i, K_i, P_i \tilde{B}_i^c)$ are all collected from Algorithm 3. Therefore, Algorithm 3 has advantages in stimulating the system and mining data for further use in next subsection.

3.3 | Optimal robust collision-free formation output control

Since multi-vehicle systems with physical sizes usually complete collaborative tasks in complex environments, obstacle avoidance is another requirement that should be considered in robust formation control. In essence, the optimal collision-free formation problem in Problem 2 is to find the optimal decision-making strategy for each agent under information constraints.

Our aim is to minimize the objective function in (8) with nonquadratic term $F(\tilde{\delta}_i, \Omega_i)$. Detection range Γ_i and safety range Π_i are considered as

$$\Gamma_i = \{y_i | r_i + r_{ob} \leq \|y_i - o_j\| \leq R_i + r_{ob}\}, \quad \Pi_i = \{y_i | \|y_i - o_j\| \geq R_i + r_{ob}\},$$

where R_i and r_i are the reaction range and size of agent i , r_{ob} is the size of the colliding object $o_j \in \Omega_i$, where $j = 1, 2, \dots, l$ and Ω_i includes other agents and obstacles in the environment.

Taking from our previous work,¹⁹ the penalty of collision avoidance for agent i is defined as

$$h_i = \sum_{i=1}^n \sum_{j=1}^{n+l-1} \begin{cases} e^{\left(\frac{(R_i+r_{ob})^2 - \|y_i-o_j\|^2}{\|y_i-o_j\|^2 - (r_i+r_{ob})^2}\right)^2} - 1, & y_i \in \Gamma_i, \\ 0, & y_i \in \Pi_i. \end{cases}$$

The cost function in (8) is designed as

$$F(\tilde{\delta}_i, \Omega_i) = \left(h'_i (\tilde{B}_i^\delta R_i^{-1} P_i - \tilde{A}_i^\delta) \tilde{\delta}_i + \frac{1}{4} h'_i \tilde{B}_i^\delta R_i^{-1} (\tilde{B}_i^\delta)^T h'_i \right), \tag{41}$$

where h'_i represents the derivation of h_i .

The optimal control input can be designed as

$$u_i = -K_i \delta_i - \frac{1}{2} K_i P_i^{-1} h'_i, \tag{42}$$

where K_i and P_i are the solutions of Algorithm 3.

The third and final result of this article is given as

Theorem 3. Consider systems (1)–(4) satisfying Assumptions 1–4. The optimal robust collision-free formation output control problem defined in Problem 2 is solved by the controller in (42) based on the adaptive observers in (11) and Algorithms 2 and 3.

Proof. Based on Theorem 3 in our previous research,¹⁹ if P_i^* is the optimal solution of the generalized algebraic Riccati equation (39), the model-based control input to minimize the nonquadratic objective function J in (8) with $F(\tilde{\delta}_i, \Omega_i)$ in (41) can be designed as

$$\tilde{u}_i^* = -R_i^{-1} (\tilde{B}_i^\delta)^T P_i^* \tilde{\delta}_i - \frac{1}{2} R_i^{-1} (\tilde{B}_i^\delta)^T h'_i. \tag{43}$$

Based on (38), the optimal input is

$$u_i^* = -R_i^{-1} (\tilde{B}_i^\delta)^T P_i^* \delta_i - \frac{1}{2} R_i^{-1} (\tilde{B}_i^\delta)^T h'_i, \tag{44}$$

where uncertain matrices $(\tilde{A}_i^\delta, \tilde{B}_i^\delta)$ are used for solving P_i^* and u_i^* . However, the uncertainties models are usually unknown.

From Theorem 2, we can see that K_i and P_i from Algorithm 3 converges to solutions K_i^* and P_i^* of (39). Since $K_i^* = R_i^{-1} (\tilde{B}_i^\delta)^T P_i^*$, it indicates $K_i P_i^{-1}$ from Algorithm 3 converges to unknown information $R_i^{-1} (\tilde{B}_i^\delta)^T$. Therefore, the controller output u_i in (42) converges to u_i^* in (44), and Problem 2 is solved by controller (42) based on Algorithm 3 and observer in (11) by Algorithm 2. This completes the proof. ■

Remark 7. Note that the optimal control input in (42) can be regarded as a combination of two terms, formation regulator control and collision avoidance control. The former decouples the heterogeneous system and obtains the optimal formation tracking control policy, while the latter realizes the optimal formation decision under dynamic obstacle avoidance.

Remark 8. Compared with the traditional RL-based algorithms involving a discount factor γ for tracking non-asymptotically stable reference systems,^{35,42} the optimal robust formation output control Problem 1 is solved by the

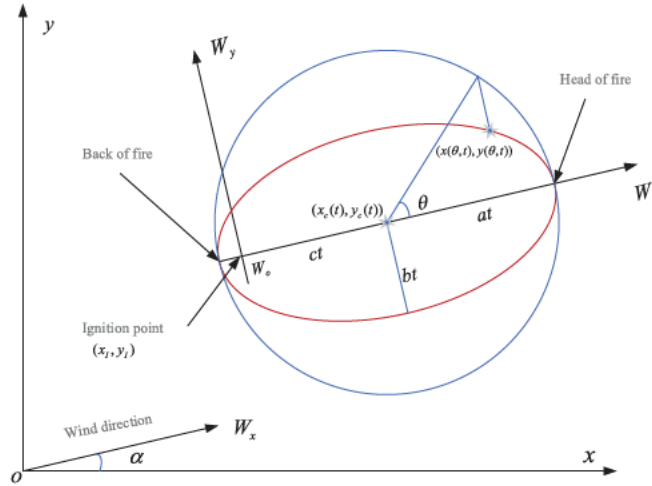


FIGURE 2 Bushfires spread model

proposed algorithms without requiring the discount factor. Although the discount factor γ is also removed in cooperative optimal output regulation solutions in Reference 41, the communication couple information is still needed in the design of the RL-based algorithms. It is worth noting that the observer and the optimization algorithms designed in Theorems 1–3 are fully distributed and model-free, such that they decouple topology from system dynamics and do not rely on any global or model information. Furthermore, differing from data-driven RL algorithms in References³⁷ and 49 with quadratic objective functions, Theorem 3 gives the optimal solution of the nonquadratic optimization to achieve the collision-free formation for heterogeneous MASs.

4 | VERIFICATION

In this section, an example of the reference formation design is given for bushfire edge tracking and patrolling tasks. Then an uncertain heterogeneous MAS consisting of four UAVs and one UGV is considered to perform the tasks by the proposed formation control strategy.

The occurrence of multiple bushfires is unfortunately not uncommon in Australia. It has been observed that the shape of the fire is approximately an ovoid, which becomes more like an ellipse during the spread of the bushfires.^{50,51}

The ellipse model based on the main wind detection is shown in Figure 2. The position of the bushfire range after time t from an ignition point is described as

$$\begin{aligned} x(\theta, t) &= ct + at \cos \theta, \\ y(\theta, t) &= bt \sin \theta, \quad 0 \leq \theta < 2\pi, \end{aligned} \tag{45}$$

where (W_x, W_o, W_y) is the wind coordinate system from the origin (ignition point), and (x, o, y) denotes the inertial coordinate system. Coordinate (x_c, y_c) indicates the center of fire, and c is the moving rate of fire center in the direction of wind. Parameters a and b denote the linear fire spread rates along the wind direction and vertical wind direction, respectively. The angle of wind speed with respect to the x -axis of the inertial system is α . When $0 \leq \theta < 2\pi$, the position set $(x(\theta, t), y(\theta, t))$ represents the edge of the bushfires. The length of the major axes at and minor axes bt of the ellipse increases with time.

To perform the tracking task, a virtual leader is designed as the moving center of the ellipse. Define the state and output variables of the virtual leader as $x_0 = \text{col}(x_c, y_c, \dot{x}_c, \dot{y}_c)$ and $y_0 = \text{col}(x_c, y_c)$, respectively. Based on the description of model (45), the state and output matrices are obtained as

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{46}$$

In the wind coordinate system (W_x, W_o, W_y) , a TVF based on the fire spread model in (45) is designed as follows

$$\begin{aligned} f_i^x &= at \cos (wt + \theta_i^0), \\ f_i^y &= bt \sin (wt + \theta_i^0), \\ \dot{f}_i^x &= a \cos (wt + \theta_i^0) - at\omega \sin (wt + \theta_i^0), \\ \dot{f}_i^y &= b \sin (wt + \theta_i^0) + bt\omega \cos (wt + \theta_i^0), \end{aligned} \tag{47}$$

where (f_i^x, f_i^y) represents the positions of i th points, which are evenly distributed over an ellipse, moving along the edge of the ellipse at an angular rate ω . The initial patrol angle is θ_i^0 .

Hence, the dynamics of the TVF in (47) satisfies

$$\begin{bmatrix} f_i^x \\ f_i^y \\ \dot{f}_i^x \\ \dot{f}_i^y \end{bmatrix}_W = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \omega^2 & 0 & 0 & -\frac{2a\omega}{b} \\ 0 & \omega^2 & \frac{2b\omega}{a} & 0 \end{bmatrix} \begin{bmatrix} f_i^x \\ f_i^y \\ \dot{f}_i^x \\ \dot{f}_i^y \end{bmatrix}_W, \tag{48}$$

where the subscript W denotes that the vector is in the wind coordinate system. Similarly, the subscript I represents that the vector is in the inertial coordinate system.

Note that model (45) is presented in the wind coordinate system (W_x, W_o, W_y) . Using the following direction cosine matrices (DCM)⁵² between the wind coordinate system (W_x, W_o, W_y) and the inertial coordinate system (x, o, y) yields,

$$DCM_{W2I} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 & 0 \\ \sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & \cos \alpha & -\sin \alpha \\ 0 & 0 & \sin \alpha & \cos \alpha \end{bmatrix},$$

and $DCM_{I2W} = DCM_{W2I}^{-1} = DCM_{W2I}^T$.

The TVF model in (47) is transformed to the inertial coordinate system as

$$\begin{bmatrix} f_i^x \\ f_i^y \\ \dot{f}_i^x \\ \dot{f}_i^y \end{bmatrix}_I = DCM_{W2I} A_w^f DCM_{I2W} \begin{bmatrix} f_i^x \\ f_i^y \\ \dot{f}_i^x \\ \dot{f}_i^y \end{bmatrix}_W, \quad A_w^f = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \omega^2 & 0 & 0 & -\frac{2a\omega}{b} \\ 0 & \omega^2 & \frac{2b\omega}{a} & 0 \end{bmatrix}.$$

If the state in (4) is designed as $f_i = [f_i^x, f_i^y, \dot{f}_i^x, \dot{f}_i^y]^T_I$, the system matrices in the inertial coordinate system become

$$A_i^f = DCM_{W2I} A_w^f DCM_{I2W}, \quad C_i^f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{49}$$

A heterogeneous MAS is considered with two Qball-X4 UAVs,⁵³ two Qball2 UAVs,¹⁹ and one Qbot2 UGV.⁹ Based on the MAS models,⁹ the latitudinal dynamics of UAVs and dynamics of UGV can be simplified to the following linear fourth-order system and second-order system, respectively.

$$\begin{aligned} \begin{bmatrix} p_i^x \\ v_i^x \\ \theta_i^x \\ q_i \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^x \\ v_i^x \\ \theta_i^x \\ q_i \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_i^m L_i}{I_i} \end{bmatrix} u_i^x + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \omega^x, & \begin{bmatrix} p_i^y \\ v_i^y \\ \phi_i^y \\ p_i \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^y \\ v_i^y \\ \phi_i^y \\ p_i \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_i^m L_i}{I_i} \end{bmatrix} u_i^y + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \omega^y, \text{ for } i = 1, 2, 3, 4, \\ &= A_i^x x_i^x + B_i^x u_i^x + D_i^x \omega^x, & & = A_i^y x_i^y + B_i^y u_i^y + D_i^y \omega^y, \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} \dot{p}_i^x \\ \dot{v}_i^x \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^x \\ v_i^x \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_i^x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \omega^x, & \begin{bmatrix} \dot{p}_i^y \\ \dot{v}_i^y \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_i^y \\ v_i^y \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_i^y + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \omega^y, \text{ for } i = 5, \\ &= A_i^x x_i^x + B_i^x u_i^x + D_i^x \omega^x, & &= A_i^y x_i^y + B_i^y u_i^y + D_i^y \omega^y, \end{aligned}$$

where g is the gravity acceleration. For the i th agent, (p_i^x, p_i^y) and (v_i^x, v_i^y) represent the horizontal positions and velocities in the X - and Y -directions, respectively. For UAVs, (ϕ_i^x, θ_i^x) and (p_i, q_i) denote the roll angle, the pitch angle, roll angular rate, and pitch angular rate respectively. k_i^m and L_i are force coefficient and the lever arm, respectively, I_i^x and I_i^y indicate moments of inertia in the X - and Y -directions.

In terms of heterogeneous MAS (1), the system states, outputs and matrices are

$$\begin{aligned} x_i &= \text{col}(x_i^x, x_i^y) = \text{col}(p_i^x, v_i^x, p_i^y, v_i^y, \phi_i^x, p_i, \theta_i^x, q_i, p_i^y, v_i^y, \phi_i^y, p_i), \text{ for } i = 1, 2, 3, 4, \\ x_i &= \text{col}(x_i^x, x_i^y) = \text{col}(p_i^x, v_i^x, p_i^y, v_i^y), \text{ for } i = 5, \quad y_i = \text{col}(p_i^x, p_i^y), \text{ for } i = 1, 2, 3, 4, 5, \\ \bar{A}_i &= \text{blkdiag}(\bar{A}_i^x, \bar{A}_i^y) = \text{blkdiag}(A_i^x + \Delta A_i^x, A_i^y + \Delta A_i^y), \text{ for } i = 1, 2, 3, 4, 5, \\ \bar{B}_i &= \text{blkdiag}(\bar{B}_i^x, \bar{B}_i^y) = \text{blkdiag}(B_i^x + \Delta B_i^x, B_i^y + \Delta B_i^y), \text{ for } i = 1, 2, 3, 4, 5, \\ \bar{D}_i &= \text{blkdiag}(\bar{D}_i^x, \bar{D}_i^y) = \text{blkdiag}(D_i^x + \Delta D_i^x, D_i^y + \Delta D_i^y), \text{ for } i = 1, 2, 3, 4, 5. \end{aligned}$$

For $i = 1, 2$, $k_i^m = 12$ N, $L_i = 0.2$ m, and $I_i^x = I_i^y = 0.03$ kg m², while for $i = 3, 4$, $k_i^m = 120$ N, $L_i = 0.2$ m, and $I_i^x = I_i^y = 0.03$ kg m². Due to load changes and the unmodeled dynamics, the following system uncertainties are considered in this MAS.

$$\begin{aligned} \Delta A_i^k &= \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \Delta B_i^k = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \Delta D_i^k = \begin{bmatrix} 0 \\ 0 \\ 0.5 \\ 0 \end{bmatrix}, \text{ } i = 1, 2, \text{ } k = x, y, \\ \Delta A_i^k &= \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \Delta B_i^k = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \Delta D_i^k = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.5 \end{bmatrix}, \text{ } i = 3, 4, \text{ } k = x, y, \\ \Delta A_i^k &= \begin{bmatrix} 0 & 0.1 \\ 0 & 0.1 \end{bmatrix}, \Delta B_i^k = \begin{bmatrix} 0 \\ 0.2 \end{bmatrix}, \Delta D_i^k = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \text{ } i = 5, \text{ } k = x, y. \end{aligned}$$

For the bushfire edge tracking and patrolling tasks, the disturbance is assumed to be the main wind disturbance with $A_w = 0$, $w_0 = [w_0^x, w_0^y]^T = [2, 1]$ in (2). The dynamics of the fire center regarding as virtue leader is given in (46) with the initial state $x_0 = [0 \ 1 \ 2 \ 1]^T$. The parameters in the design of the TVF are chosen as $a = 0.2$, $b = 0.1$, and $\omega = 0.8$. For the TVF of UGV, the intention of setting patrolling angular rate $\omega = 0$ is to make the UGV only track the back of the bushfire instead of patrolling on its edge. The initial patrolling angle of the TVF is $\theta^0 = (0, \pi/2, \pi, 3\pi/2, 0)$. The obstacle is located at (17, 17) m with a radius of 1 m. The radius of UAV and UGV are 0.5 and 0.3 m, respectively. Note that the dynamics of UAVs in Z -direction is not considered here. The aircraft is assumed to fly in a plane with an altitude of 20 m.

The communication topology of the MAS with a virtual leader is shown in Figure 3. The parameters of distributed observer (11) are set as $c = 20$ and $d = 1$. To verify the convergence of Algorithms 1–3, the parameters in the Riccati equations (21) and (39) are set as $R_i = 2I$, $Q = Q_i = 0.1I$. The iteration updates by least square technique in every 10 samples, each sampling interval is 0.1 s.

The simulation results of the observation and the formation control based on Theorems 1–3 are shown in Figures 4–9. Table 1 lists the comparison of the three RL algorithms, and the convergence trends of each algorithm are displayed in Figures 4, 5, and 7, respectively. Based on partial model information, only P_{oi} is solved by Algorithm 1 for observer (11), and the norm errors between the optimal solution P_0^* of the Riccati equation (18) and obtained matrix P_{oi} of the five

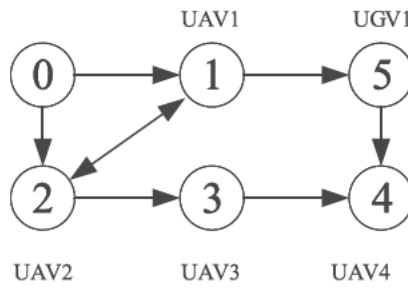


FIGURE 3 Network topology

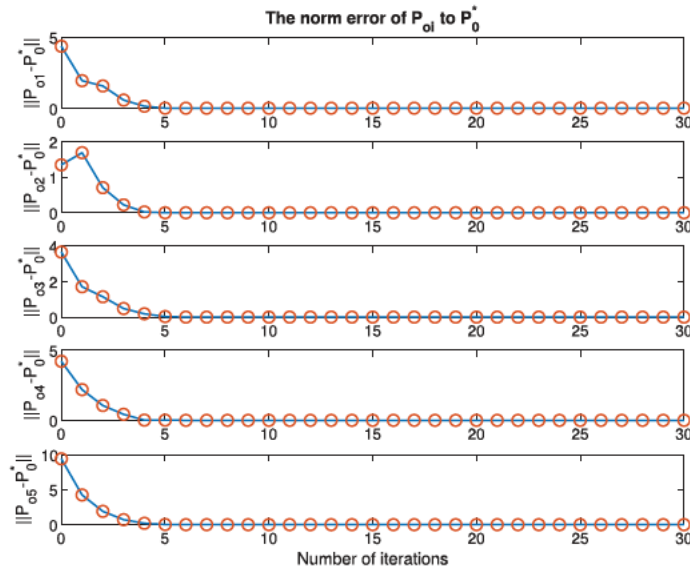


FIGURE 4 Convergence trend of P_{oi} based on Algorithm 1

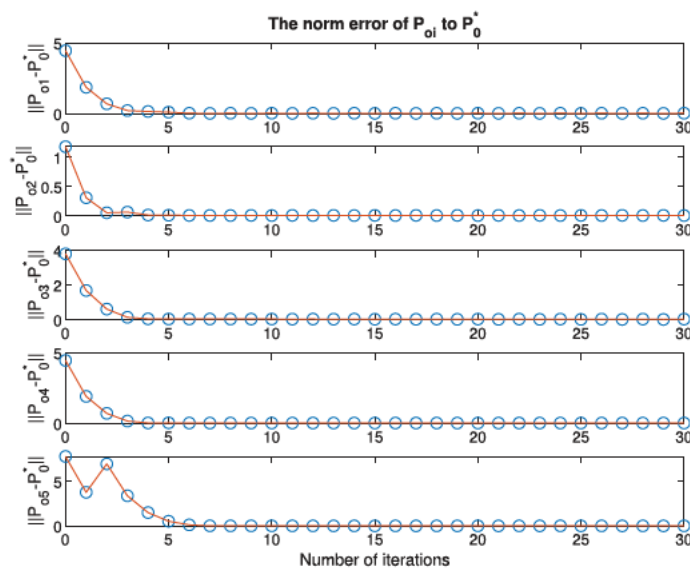


FIGURE 5 Convergence trend of P_{oi} based on Algorithm 2

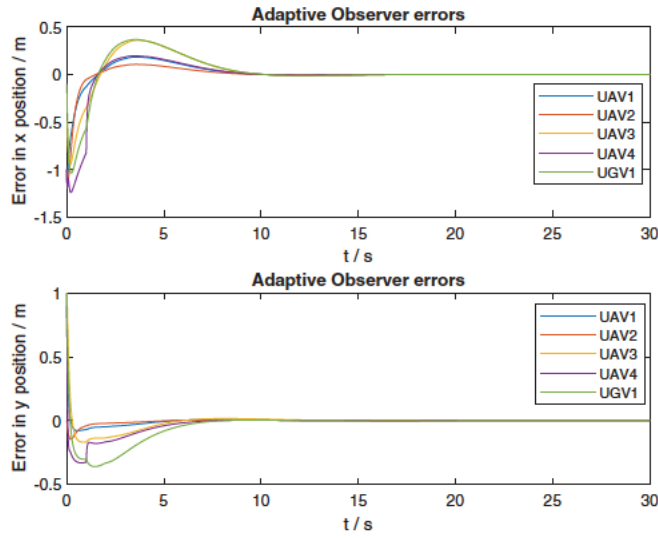


FIGURE 6 Observation errors under Algorithm 2

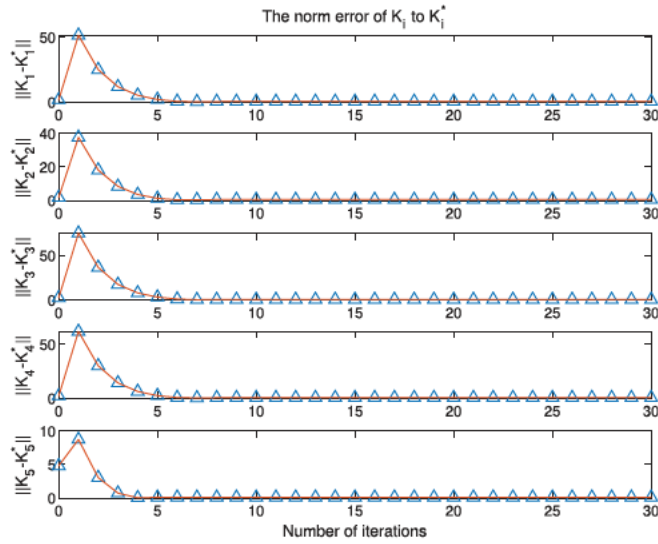


FIGURE 7 Convergence trend of K_i based on Algorithm 3

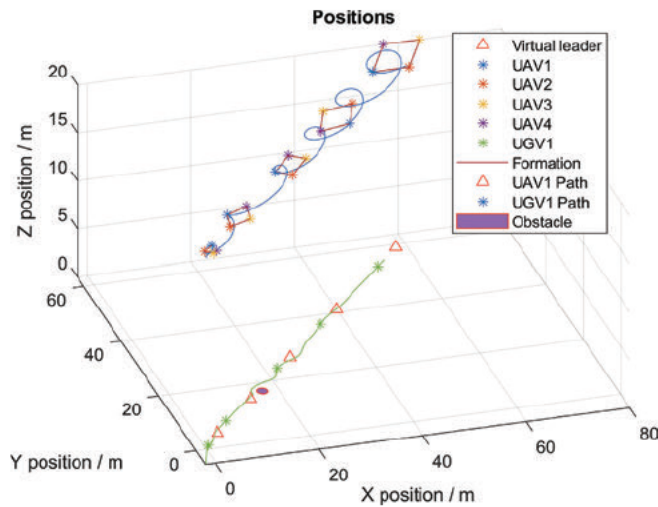


FIGURE 8 Positions of heterogeneous MASs under Algorithms 2 and 3

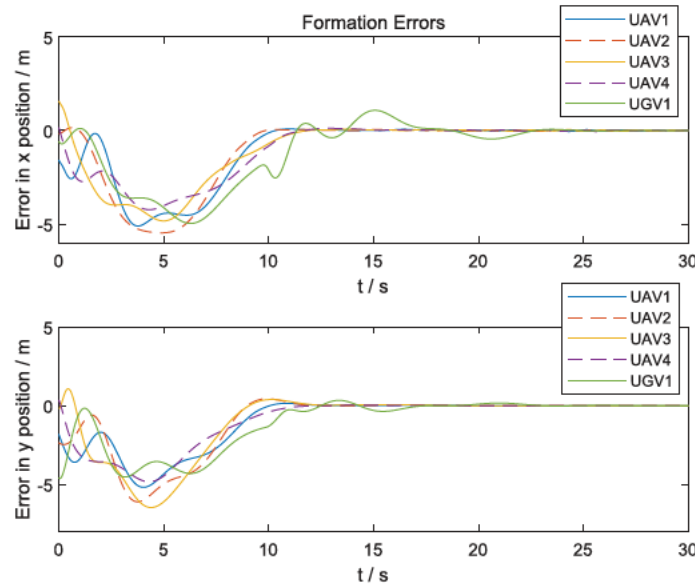


FIGURE 9 Formation errors output under Algorithms 2 and 3

TABLE 1 Comparison of Algorithms 1–3

RL algorithms	Feature	Parameters to be solved	Convergence steps of iterations				
			Agent 1	Agent 2	Agent 3	Agent 4	Agent 5
Algorithm 1	Observed model-based	P_{oi}	5	4	5	4	5
Algorithm 2	Model-free	(P_{oi}, K_{oi})	6	5	5	5	7
Algorithm 3	Model-free	$(P_i, K_i, P_i \tilde{B}_i^c)$	7	6	7	7	5

agents converge to within the threshold after 5, 4, 5, 4, and 5 iterations, respectively. By adding noise term e_{di} to stimulate systems, both solutions of P_{oi} and K_{oi} are collected by Algorithm 1 for observer (11) without requiring model information. From Figure 5, the convergence steps of iterations for each agent are 6, 5, 5, 5, and 7, respectively. Similarly, Figure 7 illustrates the norm error between the solution K_i^* solved by the Riccati equation (39) and the matrix K_i in model-free Algorithm 3, where three parameters are solved by the least squares technique. The norm errors converged within the threshold after 7, 6, 7, 7, and 5 iterations, respectively. More iterations are needed to solve more parameters for model-free algorithms, compared to the model-based algorithm. Unlike the existing works on the learning-based output regulation,⁵⁴ no discount factor is required in the developed algorithms to track the non-asymptotically stable reference systems. Note that the convergence rates of all policy iteration algorithms are acceptable, which are faster than the traditional value iteration algorithms.⁴¹ Due to the off-policy updating strategy, noise item not only has no effect on the convergence of the solutions, but also enables the PE condition to be satisfied while stimulating the system.

The observer errors of each agent based on Algorithm 2 is shown in Figure 6. Observation errors for exosystem converge to 0 within 15 s. The positions of agents are drawn every 6 s in Figure 8. The trajectories of UAV1 and UGV1 are also plotted. Based on the adaptive observers in (11) and the formation controller presented in (42), the formation output errors of heterogeneous MAS (1) with uncertainties and disturbance converge to zero within 30 s, as shown in Figure 9. During the movements of all agents, the closest distance is 2.15 m between UGV1 and the obstacle, which is greater than the safe distance of 1.3 m. It shows that the behaviors of agents match the desired tracking and patrolling reference formations, and the agents can avoid collisions.

5 | CONCLUSION

In this article, a learning-based robust optimal formation control strategy is proposed for unified heterogeneous MASs subject to uncertainties, disturbance, and complex environments. Only using local information, a distributed

observer is developed to learn exosystem information by the adaptive control and the off-policy RL method. To solve the optimal robust collision-free formation problem, a model-free RL algorithm is proposed to minimize the non-quadratic objective function, without the need for any model information. The results obtained from the simulation study verify the effectiveness and robustness of the proposed distributed observer and robust formation strategies. Our future work will consider the switching nonlinear virtual leader corresponding to dynamic collaborative tasks for unknown nonlinear heterogeneous MASs under attacks based on resilient control⁵⁵ and deep reinforcement learning.⁵⁶

CONFLICT OF INTEREST

The authors declared that they have no potential conflict of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Peng Shi  <https://orcid.org/0000-0001-8218-586X>

REFERENCES

1. Li X, Tang Y, Karimi HR. Consensus of multi-agent systems via fully distributed event-triggered control. *Automatica*. 2020;116:108898.
2. Shi P, Yan B. A survey on intelligent control for multiagent systems. *IEEE Trans Syst Man Cybern Syst*. 2021;51(1):161-175.
3. Sun Y, Shi P, Lim CC. Event-triggered adaptive leaderless consensus control for nonlinear multi-agent systems with unknown backlash-like hysteresis. *Int J Robust Nonlinear Control*. 2021;31:7409-7424.
4. Oh KK, Park MC, Ahn HS. A survey of multi-agent formation control. *Automatica*. 2015;53:424-440.
5. Hua Y, Dong X, Li Q, Ren Z. Distributed adaptive formation tracking for heterogeneous multiagent systems with multiple nonidentical leaders and without well-informed follower. *Int J Robust Nonlinear Control*. 2020;30(6):2131-2151.
6. Wang L, Xi J, He M, Liu G. Robust time-varying formation design for multiagent systems with disturbances: extended-state-observer method. *Int J Robust Nonlinear Control*. 2020;30(7):2796-2808.
7. Dong X, Hu G. Time-varying formation control for general linear multi-agent systems with switching directed topologies. *Automatica*. 2016;73:47-55.
8. Zhu Y, Zhou L, Zheng Y, Liu J, Chen S. Sampled-data based resilient consensus of heterogeneous multiagent systems. *Int J Robust Nonlinear Control*. 2020;30(17):7370-7381.
9. Yan B, Wu C, Shi P. Formation consensus for discrete-time heterogeneous multi-agent systems with link failures and actuator/sensor faults. *J Franklin Inst*. 2019;356(12):6547-6570.
10. Li X, Shi P, Wang Y. Distributed cooperative adaptive tracking control for heterogeneous systems with hybrid nonlinear dynamics. *Nonlinear Dyn*. 2019;95(3):2131-2141.
11. Su Y, Huang J. Cooperative output regulation of linear multi-agent systems. *IEEE Trans Automat Contr*. 2012;57(4):1062-1066.
12. Huang J. *Nonlinear Output Regulation: Theory and Applications*. SIAM; 2004.
13. Yang R, Zhang H, Feng G, Yan H, Wang Z. Robust cooperative output regulation of multi-agent systems via adaptive event-triggered control. *Automatica*. 2019;102:129-136.
14. Xu Z, Ni H, Karimi HR, Zhang D. A Markovian jump system approach to consensus of heterogeneous multiagent systems with partially unknown and uncertain attack strategies. *Int J Robust Nonlinear Control*. 2020;30(7):3039-3053.
15. Chen BS, Wang CP, Lee MY. Stochastic robust team tracking control of multi-UAV networked system under Wiener and Poisson random fluctuations. *IEEE Trans Cybern*. 2020;1-14. doi:10.1109/TCYB.2019.2960104
16. Lin H, Chen K, Lin R. Finite-time formation control of unmanned vehicles using nonlinear sliding mode control with disturbances. *Int J Innovat Comput Inf Control*. 2019;16(6):2341-2353.
17. Yu H, Shi P, Lim CC, Wang D. Formation control for multi-robot systems with collision avoidance. *Int J Control*. 2019;92(10):2223-2234.
18. Liu Y, Yu H, Shi P, Lim CC. Formation control and collision avoidance for a class of multi-agent systems. *J Franklin Inst*. 2019;356(10):5395-5420.
19. Yan B, Shi P, Lim CC, Wu C, Shi Z. Optimally distributed formation control with obstacle avoidance for mixed-order multi-agent systems under switching topologies. *IET Control Theory Appl*. 2018;12(13):1853-1863.
20. Guo Y, Zhou J, Li G, Zhang J. Robust formation tracking and collision avoidance for uncertain nonlinear multi-agent systems subjected to heterogeneous communication delays. *Neurocomputing*. 2020;395:107-116.
21. Jiang W, Wen G, Peng Z, Huang T, Rahmani RA. Fully distributed formation-containment control of heterogeneous linear multiagent systems. *IEEE Trans Automat Contr*. 2019;64(9):3889-3896.

22. Cai H, Huang J. Output based adaptive distributed output observer for leader–follower multiagent systems. *Automatica*. 2021;125:109413.
23. Mac TT, Copot C, Tran DT, De Keyser R. Heuristic approaches in robot path planning: a survey. *Robot Auton Syst*. 2016;86:13-28.
24. Rasekhipour Y, Khajepour A, Chen SK, Litkouhi B. A potential field-based model predictive path-planning controller for autonomous road vehicles. *IEEE Trans Intell Transp Syst*. 2017;18(5):1255-1267.
25. Zhang XH, Han GS, Guan ZH, Li J, Zhang DX, Liao RQ. Robust multi-tracking of heterogeneous multi-agent systems with uncertain nonlinearities and disturbances. *J Franklin Inst*. 2018;355(8):3677-3690.
26. Wang YW, Liu XK, Xiao JW, Shen Y. Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control. *Automatica*. 2018;93:26-32.
27. Bardi M, Capuzzo-Dolcetta I. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. SIAM; 2008.
28. Gao Z, Yang P, Qian M. Iterative learning observer-based fault tolerant control approach for satellite attitude system with mixed actuator faults. *ICIC Express Lett*. 2019;13:635-643.
29. Deng C, Che WW, Shi P. Cooperative fault-tolerant output regulation for multiagent systems by distributed learning control approach. *IEEE Trans Neural Netw Learn Syst*. 2020;31(11):4831-4841. doi:10.1109/TNNLS.2019.2958151
30. Zhao W, Liu H, Lewis FL. Robust formation control for cooperative underactuated quadrotors via reinforcement learning. *IEEE Trans Neural Netw Learn Syst*. 2020;1-11. doi:10.1109/TNNLS.2020.3023711
31. Kiumarsi B, Vamvoudakis KG, Modares H, Lewis FL. Optimal and autonomous control using reinforcement learning: a survey. *IEEE Trans Neural Netw Learn Syst*. 2018;29(6):2042-2062.
32. Ryoma Watanuki TH, Aodai T. Vision-based behavior acquisition by deep reinforcement learning in multi-robot environment. *ICIC Express Lett B Appl*. 2020;11(3):237-244.
33. Liu M, Wan Y, Lewis FL, Lopez VG. Adaptive optimal control for stochastic multiplayer differential games using on-policy and off-policy reinforcement learning. *IEEE Trans Neural Netw Learn Syst*. 2020;31(12):5522-5533.
34. Werbos PJ. A menu of designs for reinforcement learning over time. In: Werbos PJ, Miller W, Sutton R, eds. *Neural Networks for Control*. MIT Press; 1990:67-95.
35. Modares H, Lewis FL, Jiang ZP. H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning. *IEEE Trans Neural Netw Learn Syst*. 2015;26(10):2550-2562.
36. Liu H, Peng F, Modares H, Kiumarsi B. Heterogeneous formation control of multiple rotorcrafts with unknown dynamics by reinforcement learning. *Inf Sci*. 2021;558:194-207.
37. Odekunle A, Gao W, Davari M, Jiang ZP. Reinforcement learning and non-zero-sum game output regulation for multi-player linear uncertain systems. *Automatica*. 2020;112:108672.
38. Jiang Y, Kiumarsi B, Fan J, Chai T, Li J, Lewis FL. Optimal output regulation of linear discrete-time systems with unknown dynamics using reinforcement learning. *IEEE Trans Cybern*. 2020;50(7):3147-3156.
39. Yang Y, Modares H, Wunsch DC, Yin Y. Optimal containment control of unknown heterogeneous systems with active leaders. *IEEE Trans Control Syst Technol*. 2019;27(3):1228-1236.
40. Yang Y, Modares H, Wunsch DC, Yin Y. Leader–follower output synchronization of linear heterogeneous systems with active leader using reinforcement learning. *IEEE Trans Neural Netw Learn Syst*. 2018;29(6):2139-2153.
41. Gao W, Mynuddin M, Wunsch DC, Jiang ZP. Reinforcement learning-based cooperative optimal output regulation via distributed adaptive internal model. *IEEE Trans Neural Netw Learn Syst*. 2021;1-12. doi:10.1109/TNNLS.2021.3069728
42. Modares H, Nageshrao SP, Lopes GAD, Babuška R, Lewis FL. Optimal model-free output synchronization of heterogeneous systems using off-policy reinforcement learning. *Automatica*. 2016;71:334-341.
43. Qu Z. *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer-Verlag; 2009.
44. Vrabie D, Pastravanu O, Abu-Khalaf M, Lewis FL. Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica*. 2009;45(2):477-484.
45. Modares H, Lewis FL. Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning. *IEEE Trans Automat Contr*. 2014;59(11):3051-3056.
46. Cai H, Lewis FL, Hu G, Huang J. The adaptive distributed observer approach to the cooperative output regulation of linear multi-agent systems. *Automatica*. 2017;75:299-305.
47. Jiang Y, Jiang ZP. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. *Automatica*. 2012;48(10):2699-2704.
48. Zhang H, Chen J, Wang Z, Fu C, Song S. Distributed event-triggered control for cooperative output regulation of multiagent systems with an online estimation algorithm. *IEEE Trans Cybern*. 2020. doi:10.1109/TCYB.2020.2991761
49. Gao W, Liu Y, Odekunle A, Yu Y, Lu P. Adaptive dynamic programming and cooperative output regulation of discrete-time multi-agent systems. *Int J Control Autom Syst*. 2018;16(5):2273-2281.
50. Perry G. Current approaches to modelling the spread of wildland fire: a review. *Prog Phys Geogr*. 1998;22(2):222-245.
51. Yan B, Shi P, Lim CC. Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy. *IEEE Trans Autom Sci Eng*. 2021;1-13. doi:10.1109/TASE.2021.3103877

52. Valavanis KP, Vachtsevanos GJ. *Handbook of Unmanned Aerial Vehicles*. Springer; 2015.
53. Liu Z, Yuan C, Zhang Y, Luo J. A learning-based fault tolerant tracking control of an unmanned quadrotor helicopter. *J Intell Robot Syst*. 2016;84:145-162.
54. Yaghmaie FA, Gunnarsson S, Lewis FL. Output regulation of unknown linear systems using average cost reinforcement learning. *Automatica*. 2019;110:108549.
55. Ma R, Shi P, Wang Z, Wu L. Resilient filtering for cyber-physical systems under denial-of-service attacks. *Int J Robust Nonlinear Control*. 2020;30(5):1754-1769.
56. Kiran BR, Sobh I, Talpaert V, et al. Deep reinforcement learning for autonomous driving: a survey. *IEEE Trans Intell Transp Syst*. 2021;1-18. doi:10.1109/TITS.2021.3054625

How to cite this article: Yan B, Shi P, Lim C-C, Shi Z. Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning. *Int J Robust Nonlinear Control*. 2022;32(5):2683-2704. doi: 10.1002/rnc.5828

Chapter 6

Event and Learning-based Resilient Formation Control for Multi-agent Systems under DoS Attacks

Statement of Authorship

Title of Paper	Event and learning-based resilient formation control for multi-agent systems under DoS attacks
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	B. Yan, Y. Sun, P. Shi and C. -C. Lim, "Event and learning-based resilient formation control for multi-agent systems under DoS attacks," submitted and under review.

Principal Author

Name of Principal Author (Candidate)	Bing Yan		
Contribution to the Paper	Conceptualization, methodology, experiments, validation and writing-original draft		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25 Jul 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Yuan Sun		
Contribution to the Paper	Experiments and writing-review		
Signature		Date	25 Jul 2022

Name of Co-Author	Peng Shi		
Contribution to the Paper	Polishing, checking and verification.		
Signature		Date	25 Jul 2022

Name of Co-Author	Cheng-Chew Lim		
Contribution to the Paper	Review, refine and validate		
Signature		Date	26/7/22

6.1 Introduction

Considering denial-of-service (DoS) attacks, this chapter presents a novel resilient and robust two-layer controller with a brand-new reinforcement learning (RL) condition to address TVF problems for unknown heterogeneous MAS. The design is distributed and model-free at the cyber-layer and the physical system layer. An event-based resilient observer is provided at the cyber-layer to remove global information of communication and deal with DoS attacks. The communication load can be reduced under attacks and the Zeno behavior can be avoided. In the physical system layer, an RL rank condition for the TVF controller is developed for unknown heterogeneous MAS. Compared with the RL algorithms in Chapter 5, the new rank condition can automatically adjust online data collection time, thereby improving online learning and optimization performance. Experiments of multi-UGV area scanning formations are conducted. The comparative experimental results verify the resilience of the proposed online event and learning-based control method under different parameters of DoS attacks.

6.2 Publication

B. Yan, Y. Sun, P. Shi and C. -C. Lim, "Event and learning-based resilient formation control for multi-agent systems under DoS attacks," submitted, under review.

Event and learning-based resilient formation control for multi-agent systems under DoS attacks

Bing Yan, Yuan Sun, Peng Shi*, *Fellow, IEEE*, and Cheng-Chew Lim, *Senior Member, IEEE*

Abstract

In this paper, a novel resilient time-varying formation control strategy is developed for multi-agent systems under denial-of-service (DoS) attacks based on an event-triggered observer and online reinforcement learning. The approach is distributed and model-free with a decoupled two-layer design for the cyber-layer and the physical system layer. The event-based resilient observer is proposed to estimate an exosystem under DoS attacks with dual adaptive laws to remove global information of a communication topology at the cyber-layer. In the physical system layer, an optimal formation output control is designed for multi-agent systems based on the output regulation framework and off-policy reinforcement learning with a new rank condition. Finally, experiments with unmanned ground vehicles for area scanning formations are conducted to verify the effectiveness and resilience of the proposed online event and learning-based control method.

Index Terms

Resilient event-triggered strategy; reinforcement learning-based formation; multi-agent systems; denial-of-service attacks

I. INTRODUCTION

Formation control is a way for multi-agent systems (MAS) to perform tasks towards the common goal by changing the motions of each agent and the distribution of the relative positions between agents [1]. Operating in complex communication environments, the cyber-layer between agents is threatened

Bing Yan, Yuan Sun, Peng Shi and Cheng-Chew Lim are with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia. (e-mails: bing.yan@adelaide.edu.au, yuan.sun01@adelaide.edu.au, peng.shi@adelaide.edu.au, cheng.lim@adelaide.edu.au) *Corresponding author: Peng Shi.

by malicious attacks [2], and the physical system layer is subjected to uncertain even unknown model information and local incomplete information due to limited onboard sensors [3]. The multiple constraints at the two layers impact the communication security, formation stability, and performance optimization of the MAS. Therefore, achieving resilient formation control and optimization is a significant challenge for unknown MAS.

Compared with homogeneous MAS [4], [5], heterogeneous MAS consisting of entities with different dynamics have been intensively studied due to their flexibility in completing complex tasks [6], [7]. Output regulation framework with internal model control has been proposed enabling different entities to achieve output consensus by tracking a common reference system, called an exosystem [8]. The framework has been extended to solve time-varying formation (TVF) control problems of heterogeneous MAS [9], [10]. However, most of the existing approaches focus on the stability of consensus-based formation systems or output regulation-based formation systems. Formation performance optimization has not been fully considered. Furthermore, how to design effective TVF that can execute collaborative tasks in real-world applications also demands studies.

Since the agents communicate through a shared network, cyber-security issues arising from potential malicious attacks should be considered. The denial-of-service attack (DoS) is a typical cyber-attack for MAS in which the perpetrator seeks to block communication between agents and prevent the transmission of measurement data to the controller [4]. A non-sampling output regulation approach has been provided to deal with DoS attacks based on Linear Matrix Inequalities and data-driven algorithms [11]. As each agent is commonly equipped with a certain number of devices, conserving network resources is as important as ensuring reliable communication. Event-based sampling strategies have been proved to be effective in improving communication resilience and reducing network load under DoS attacks [12], [13]. Intensive works on Zeno-free event-triggered strategies have been proposed to solve consensus and formation control problems for MAS [14], [15]. However, most of the existing resilient methods for DoS attacks are not distributed due to the use of global graph information, and the dynamic global information is generally difficult to obtain in real-world applications.

Furthermore, uncertain and even unknown system information challenges the robustness of MAS and its capability for performance optimization. Although robust formation control approaches [16] can deal with system uncertainties, it is not sufficient to solve model-free problems. Reinforcement learning (RL) has been proven to effectively achieve online optimization and remove model information through data-driven approaches [17]. For instance, off-policy RL has been developed to deal with consensus problems [18] and output regulation problems [19] without attacks, where off-policy means the target policy differs from the behavior policy in RL [19]. While RL is a viable solution, for practical use, the resilience and

the online optimization performance of RL algorithms under attack are still challenging conundrums deserving further studies.

Therefore, we present a novel event and learning-based formation control strategy in this paper, and its main contributions are threefold:

1. A novel resilient event-triggered observer with dual adaptive laws is proposed for exosystem estimation via communication networks under DoS attacks, and the Zeno behavior can be avoided. Our design ensures efficient estimation under attacks, reduces communication frequency, and removes the dependence on global communication topology information at the same time.
2. A robust and optimal formation control is provided for uncertain and unknown heterogeneous MAS based on an off-policy RL algorithm with a new rank condition on the persistence of excitation to achieve online optimization and automatically adjust online data collection time.
3. The resilient and robust strategy is a two-layer distributed and model-free solution for the time-varying formation control problem. It can be applied to practical MAS to support safe and efficient collaborations under DoS attacks when performing tasks via team formations.

II. PROBLEM FORMULATION

Consider a heterogeneous MAS of n agents, in which each agent can be described by

$$\dot{x}_i = \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{D}_i \omega, \quad y_i = \tilde{C}_i x_i, \quad i = 1, 2, \dots, n \quad (1)$$

where $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}$, $y_i \in \mathbb{R}^p$, and $\omega \in \mathbb{R}^\omega$ are the i th agent's state, input, output variables and disturbance. Matrices $\tilde{\star}_i = \star_i + \Delta\star_i$, $\star = A, B, C, D$ represent corresponding system matrices and $\Delta\star_i$ is the uncertainties.

In order to decouple the heterogeneous agent dynamics from networks, an exosystem is introduced for MAS to reach output consensus in the output regulation framework [8]. The exosystem can be regarded as a common task system generating reference signals for tracking and disturbance signals for elimination. Therefore, we define an exosystem as

$$\dot{\eta} = A^\eta \eta, \quad y^\eta = C^\eta \eta, \quad (2)$$

where $\eta = \text{col}(x_0, \omega) \in \mathbb{R}^{n_\eta}$ is the system state variable, function col denotes a column vector composed of virtual leader x_0 and disturbance ω , and $y^\eta \in \mathbb{R}^p$ is the output of exosystem. System state and output matrices are $A^\eta = \text{blkdiag}(A_0, A_\omega)$, and $C^\eta = [C_0, 0]$, where $\text{blkdiag}(A_0, A_\omega)$ denotes a block diagonal matrix created by aligning matrices A_0 and A_ω . Note that not every agent has access to the exosystem

information. For example, it is difficult to ensure that every agent in a large-scale MAS can directly obtain global task information.

To perform tasks towards the common goal, i th time-varying formation systems for i th agent are

$$\dot{f}_i = A_i^f f_i, \quad y_i^f = C_i^f f_i, \quad i = 1, 2, \dots, n \quad (3)$$

where $f_i \in \mathbb{R}^{n_f}$ is i th TVF state and $y_i^f \in \mathbb{R}^p$ is its output. Its system matrices are A_i^f and C_i^f .

System (1) with n augmented formation reference systems consisting a common exosystem and n TVF systems can be reconstructed as

$$\begin{aligned} \dot{x}_i &= \tilde{A}_i x_i + \tilde{B}_i u_i + \tilde{F}_i v_i \\ \dot{v}_i &= A_i^v v_i, \quad i = 1, 2, \dots, n \\ e_i &= \tilde{C}_i x_i - C_i^v v_i \end{aligned} \quad (4)$$

where $\tilde{F}_i = [0, \tilde{D}_i, 0]$, $A_i^v = \text{blkdiag}(A_0, A_\omega, A_i^f)$, and $C_i^v = [C_0, 0, C_i^f]$. Variable $v_i = \text{col}(x_0, \omega, f_i) \in \mathbb{R}^{n_i^v}$ is augmented state, where $n_i^v = n_\eta + n_f$, and $e_i \in \mathbb{R}^p$ denotes the formation output error.

Remark 1. Note that the system dynamics \tilde{x}_i and orders n_i can be different in heterogeneous MAS (1). A common disturbance ω is considered here for agents in their shared environment, such as the main wind disturbance. Different disturbances for different agents have been studied in our previous work [10].

We introduce the following assumptions used in the paper.

Assumption 1. *There are no eigenvalues with negative real parts in matrix A_i^v .*

Assumption 2. *The pair $(\tilde{A}_i, \tilde{B}_i)$ is controllable.*

Assumption 3. *The MAS communication graph contains a spanning tree, where the exosystem is the root node and other agent nodes are connected.*

Assumption 4. *For any $\lambda \in \sigma(A_i^v)$,*

$$\text{rank} \begin{bmatrix} A_i - \lambda I_{n_i} & B_i \\ C_i & \mathbf{0} \end{bmatrix} = n_i + p_i, \quad i = 1, 2, \dots, n$$

where A_i is the known state matrix of agent i , and A_i^v is the state matrix of the common exosystem. $\sigma(A_i^v)$ denotes the eigenvalues of A_i^v .

The DoS attacks usually act on the communication channel between agents to prevent the transmission of data. To describe the attack model [13], the c th DoS interval can be defined as

$$D_a^c = [d_c, d_c + \sigma_c), \quad c \in \mathbb{N}_0$$

where d_c and $d_c + \sigma_c$ are the off-to-on and on-to-off transitions for DoS interval D_a^c . During time interval $[s, T)$, the union of DoS intervals is

$$D_a(s, T) = \left\{ \bigcup_{c \in \mathbb{N}_0} D_a^c \right\} \cap [s, T)$$

The union of DoS inactive intervals is

$$D_w(s, T) = [s, T) \setminus D_a(s, T)$$

where \setminus indicates that working communication intervals $D_w(s, T)$ is to remove attacked intervals $D_a(s, T)$ from interval $[s, T)$.

Assumption 5. [11] *The DoS duration satisfies*

$$|D_a(s, T)| \leq \epsilon_a + \frac{T - s}{R_d} \quad (5)$$

for $T > s > 0$, where ϵ_a is positive constant, $R_d > 1$, and $|D_a(s, T)|$ is the time length of $D_a(s, T)$.

Remark 2. Note that Assumption 1 is reasonable for the exosystem and TVF with non-zero dynamics so that agent does not collide when tracking the exosystem and forming formations. Similar assumptions can be found in [10], [20]. Assumption 4 is the precondition for the output regulation to be solvable [8], which ensures that the following equations have a unique solution (Π_i^v, U_i^v)

$$\begin{aligned} \Pi_i^x A_i^v &= \tilde{A}_i \Pi_i^x + \tilde{B}_i U_i^v + \tilde{F}_i \\ \Pi_i^z A_i^v &= \Sigma_{i1} \Pi_i^z + \Sigma_{i2} (\tilde{C}_i \Pi_i^x - C_i^v) \\ 0 &= \tilde{C}_i \Pi_i^x - C_i^v \end{aligned} \quad (6)$$

where $\Pi_i^v = \text{col}(\Pi_i^x, \Pi_i^z)$, and the pair $(\Sigma_{i1}, \Sigma_{i2})$ denotes p-copy internal model pair of A_i^v [8]. Assumption 5 means that the attacked duration is not larger than a proportion of the overall interval length, and similar assumptions are used in the classical DoS attacks in [11].

The optimal resilient TVF control problem considered in this paper is to minimize the following performance index

$$\min J_i = \int_0^\infty (\zeta_i^T Q_i \zeta_i + \tilde{u}_i^T R_i \tilde{u}_i) dt, \quad i = 1, 2, \dots, n \quad (7)$$

by designing a controller for system (4) under DoS attacks, where Q_i and R_i are given positive definite matrices, $\zeta_i = \text{col}(x_i, z_i)$ is internal model-based state with compensation variable z_i . $\tilde{u}_i = u_i - U_i^v v_i$ is the input of internal model-based TVF control and error $\tilde{\zeta}_i = \zeta_i - \Pi_i^v v_i$ is equivalent to formation output error e_i , where (Π_i^v, U_i^v) is the unique solution of output regulation equation (6).

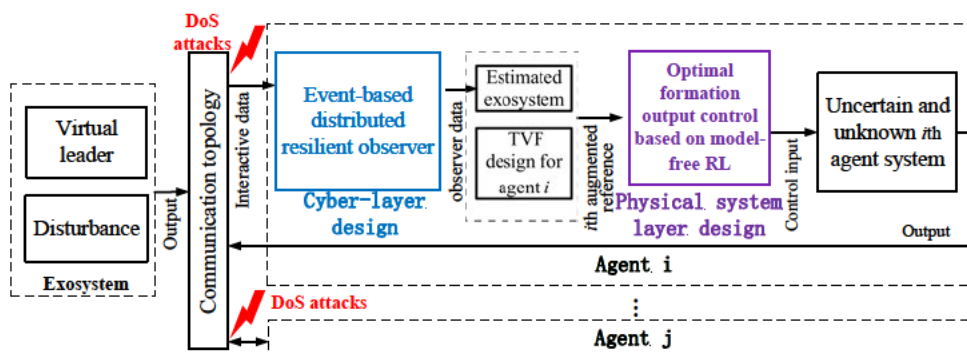


Fig. 1. Overview of the cyber-layer design and physical system layer design

III. MAIN RESULTS

In this section, a two-layer resilient formation control approach is provided to solve the optimal resilient TVF control problem in (7). The overview of the design is illustrated in Fig. 1. In the cyber-layer, an event-based distributed resilient observer is designed to estimate the exosystem based on the interactive data under DoS attacks, in which dual adaptive laws are used to remove global topology information and adjust trigger intervals to counter attacks. An optimal formation output controller is presented at the physical system layer to track augmented reference systems based on online off-policy RL without requiring model information.

A. Event-based distributed resilient observer

Since the exosystem is not reachable to each agent, they share exosystem information through communication networks. Distributed resilient observers are urgently needed to estimate shared information as the networks are subject to DoS attacks, limited network resources, and local incomplete data. An event-based observer is designed to ensure the resilience of observation and save network resources at the cyber-layer as

$$\begin{aligned}
 \dot{\hat{\eta}}_i &= A^\eta \hat{\eta}_i - K_i^\eta \theta_i \psi_i \\
 \psi_i &= \sum_{j=1}^n a_{ij} \left(e^{A^\eta(t-t_i^k)} \hat{\eta}_i(t_i^k) - e^{A^\eta(t-t_j^k)} \hat{\eta}_j(t_j^k) \right) \\
 &\quad + a_{i0} e^{A^\eta(t-t_i^k)} \left(\hat{\eta}_i(t_i^k) - \eta(t_i^k) \right) \quad t \in [t_i^k, t_i^{k+1}) \\
 \dot{\theta}_i &= \psi_i^T \psi_i, \quad \theta_i(t_0) > 1
 \end{aligned} \tag{8}$$

where $\hat{\eta}_i$ is the estimation of exosystem state η . θ_i is the first adaptive law in the observer, and $\theta_i(t_0)$ represents its initial value. Indexed exosystem as number 0, a_{i0} represents the element of the exosystem

to MAS adjacency matrix A_{n0} , and a_{ij} denotes the element of the agent to agent adjacency matrix A_{nn} .

The event-triggered instants of i th agent are represented by the time sequence $\{t_i^0, t_i^1, \dots, t_i^k, \dots \mid i = 1, 2, \dots, n, k = 0, 1, \dots\}$, and t_j^k is the j th agent's most recent triggered time. The adaptive event-based strategy is given as

$$t_i^{k+1} = \inf\{t > t_i^k \mid \phi_i \theta_i e_{\eta_i}^T e_{\eta_i} > \psi_i^T \psi_i + \gamma_1 e^{-\gamma_2 t}\} \quad (9)$$

where $\gamma_1 > 0$, $\gamma_2 > 0$, and e_{η_i} is the measurement error constructed as

$$e_{\eta_i} = e^{A^n(t-t_i^k)} \hat{\eta}_i(t_i^k) - \hat{\eta}_i, \quad t \in [t_i^k, t_i^{k+1}) \quad (10)$$

We update the second adaptive law ϕ_i in the observer by

$$\dot{\phi}_i = e_{\eta_i}^T \theta_i e_{\eta_i}, \quad \phi_i(t_0) > 1 \quad (11)$$

The observer error is defined as $\tilde{\eta}_i = \hat{\eta}_i - \eta$, where $\eta = e^{A^n(t-t_i^k)} \eta(t_i^k)$. Defining the following variables for the whole MAS as

$$\underline{\eta} = I_n \otimes \eta, \quad \tilde{\underline{\eta}} = \text{col}(\tilde{\eta}_1, \tilde{\eta}_2, \dots, \tilde{\eta}_n)$$

$$\hat{\underline{\eta}} = \text{col}(\hat{\eta}_1, \hat{\eta}_2, \dots, \hat{\eta}_n)$$

$$\Psi = \text{col}(\psi_1, \psi_2, \dots, \psi_n), \quad \Theta = \text{diag}(\theta_1, \theta_1, \dots, \theta_n)$$

$$\varphi = \text{diag}(\phi_1, \phi_2, \dots, \phi_n), \quad e_\eta = \text{col}(e_{\eta_1}, e_{\eta_2}, \dots, e_{\eta_n})$$

where \otimes represents the Kronecker product, and $\text{diag}()$ denotes a square diagonal matrix with the elements of the input vector. If the Laplacian matrix of n agents is L , the Laplacian matrix for MAS with the exosystem is

$$L_{n+1} = \begin{bmatrix} 0 & 0_{1 \times n} \\ L_0 & L_1 \end{bmatrix} \quad (12)$$

where $L_1 = L + A_{n0}$ is a nonsingular symmetric matrix under Assumption 3 [21]. From (12), we can get

$$\begin{aligned} \dot{\Psi} &= (L_1 \otimes I_{n_\eta}) \hat{\underline{\eta}} + (L_0 \otimes I_{n_\eta}) \underline{\eta} + (L_1 \otimes I_{n_\eta}) e_\eta \\ &= (L_1 \otimes I_{n_\eta}) \tilde{\underline{\eta}} + (L_1 \otimes I_{n_\eta}) e_\eta \end{aligned} \quad (13)$$

It can be obtained by taking the derivative of (8) that

$$\begin{aligned} \dot{\tilde{\underline{\eta}}} &= (I_n \otimes A^n - \Theta L_1 \otimes K_i^\eta) \tilde{\underline{\eta}} - (\Theta L_1 \otimes K_i^\eta) e_\eta \\ &= (I_n \otimes A^n) \tilde{\underline{\eta}} - (\Theta \otimes K_i^\eta) \Psi \end{aligned} \quad (14)$$

$$\dot{\Psi} = (I_n \otimes A^n) \Psi$$

$$\dot{\varphi} = e_\eta^T (\Theta \otimes I_{n_\eta}) e_\eta$$

Our first result in this paper is as follows.

Theorem 1. *Consider MAS (1) with exosystem (2) and TVF system (3) under DoS attack satisfying Assumptions 1-5. The observer error of event-based distributed resilient observer (8) with $K_i^\eta = P^{-1}$ is bounded under DoS attacks, if there is a nonsingular symmetric positive definite matrix P that satisfies the conditions (15)-(17). Furthermore, Zeno behavior can be avoided.*

$$(A^\eta)^T P + P A^\eta - I_{n_\eta} + \alpha_1 P < 0, \quad \alpha_1 > 0 \quad (15)$$

$$(A^\eta)^T P + P A^\eta - \alpha_2 P < 0, \quad \alpha_2 > 0 \quad (16)$$

$$R_d \alpha_1 > \alpha_1 + \alpha_2 \quad (17)$$

Proof. Considering the normal communication interval $t \in D_w(0, \infty)$, we define Lyapunov functions as

$$V = V_1 + V_2 + V_3 \quad (18)$$

where

$$V_1 = \tilde{\eta}^T (I_n \otimes P) \tilde{\eta} \quad (19)$$

$$V_2 = \sum_{i=1}^n \frac{1}{2} (\theta_i - c_1)^2 \quad (20)$$

$$V_3 = \sum_{i=1}^n \frac{1}{2} (\phi_i - c_2)^2 \quad (21)$$

where constants c_1 and c_2 will be designed later. From (8) and (11), it implies that

$$\theta_i > 1, \quad \phi_i > 1, \quad V > 0$$

Differentiating V_1 in (19) yields

$$\begin{aligned} \dot{V}_1 = & \tilde{\eta}^T [I_n \otimes (P A^\eta + (A^\eta)^T P) - 2\Theta L_1 \otimes I_{n_\eta}] \tilde{\eta} \\ & - 2\tilde{\eta}^T (\Theta L_1 \otimes I_{n_\eta}) e_\eta \end{aligned} \quad (22)$$

Differentiating of V_2 in (20) based on (14) leads to

$$\begin{aligned} \dot{V}_2 = & \sum_{i=1}^n (\theta_i - c_1) \dot{\theta}_i \\ = & \Psi^T ((\Theta - (c_1 - 1)I_n) \otimes I_{n_\eta}) \Psi - \Psi^T (I_n \otimes I_{n_\eta}) \Psi \\ = & \tilde{\eta}^T (L_1 (\Theta - (c_1 - 1)I_n) L_1 \otimes I_{n_\eta}) \tilde{\eta} - \Psi^T (I_n \otimes I_{n_\eta}) \Psi \\ & + 2\tilde{\eta}^T (L_1 (\Theta - (c_1 - 1)I_n) L_1 \otimes I_{n_\eta}) e_\eta \\ & + e_\eta^T (L_1 (\Theta - (c_1 - 1)I_n) L_1 \otimes I_{n_\eta}) e_\eta \end{aligned} \quad (23)$$

Taking the derivation of (21) under event-triggered schemes in (9)-(11), we obtain

$$\begin{aligned}\dot{V}_3 &= \sum_{i=1}^n (\phi_i - c_2) \dot{\phi}_i = \sum_{i=1}^n e_{\eta_i}^T (\phi_i - c_2) \theta_i e_{\eta_i} \\ &\leq \Psi^T (I_n \otimes I_{n_\eta}) \Psi - e_\eta^T (c_2 \Theta \otimes I_{n_\eta}) e_\eta + \gamma_1 e^{-\gamma_2 t}\end{aligned}\quad (24)$$

It follows from (22)-(24) that

$$\begin{aligned}\dot{V} &= \dot{V}_1 + \dot{V}_2 + \dot{V}_3 \\ &= \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P) - 2\Theta L_1 \otimes I_{n_\eta}] \tilde{\eta} \\ &\quad + \tilde{\eta}^T [L_1(\Theta - (c_1 - 1)I_n)L_1 \otimes I_{n_\eta}] \tilde{\eta} \\ &\quad - 2\tilde{\eta}^T (\Theta L_1 \otimes I_{n_\eta}) e_\eta + \gamma_1 e^{-\gamma_2 t} \\ &\quad + 2\tilde{\eta}^T (L_1(\Theta - (c_1 - 1)I_n)L_1 \otimes I_{n_\eta}) e_\eta \\ &\quad + e_\eta^T (L_1(\Theta - (c_1 - 1)I_n)L_1 \otimes I_{n_\eta} - c_2 \Theta \otimes I_{n_\eta}) e_\eta\end{aligned}\quad (25)$$

With Young's inequality, we get

$$\begin{aligned}&- 2\tilde{\eta}^T ((\Theta L_1 - L_1 \Theta L_1) \otimes I_{n_\eta}) e_\eta \\ &- 2\tilde{\eta}^T ((c_1 - 1)L_1 L_1 \otimes I_{n_\eta}) e_\eta \\ &\leq \tilde{\eta}^T ((\Theta L_1 - L_1 \Theta L_1) \otimes I_{n_\eta}) \tilde{\eta} \\ &\quad + e_\eta^T ((\Theta L_1 - L_1 \Theta L_1) \otimes I_{n_\eta}) e_\eta \\ &\quad + \tilde{\eta}^T \left(\frac{c_1 - 1}{2} L_1 L_1 \otimes I_{n_\eta} \right) \tilde{\eta} \\ &\quad + e_\eta^T (2(c_1 - 1)L_1 L_1 \otimes I_{n_\eta}) e_\eta\end{aligned}\quad (26)$$

Substituting (26) into (25) yields

$$\begin{aligned}\dot{V} &\leq \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P) \\ &\quad - (\Theta L_1 + \frac{c_1 - 1}{2} L_1 L_1) \otimes I_{n_\eta}] \tilde{\eta} + \gamma_1 e^{-\gamma_2 t} \\ &\quad + e_\eta^T [((c_1 - 1)L_1 L_1 - c_2 \Theta + L_1 \Theta) \otimes I_{n_\eta}] e_\eta\end{aligned}\quad (27)$$

Since $\Theta > I$, we obtain

$$\begin{aligned}\dot{V} &< \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P) \\ &\quad - (L_1 + \frac{c_1 - 1}{2} L_1 L_1) \otimes I_{n_\eta}] \tilde{\eta} + \gamma_1 e^{-\gamma_2 t} \\ &\quad + e_\eta^T [((c_1 - 1)L_1 L_1 - c_2 + L_1) \Theta \otimes I_{n_\eta}] e_\eta\end{aligned}\quad (28)$$

Choosing c_1 and c_2 to satisfy

$$c_1 > \frac{2(1 - \lambda_{\min}(L_1))}{\lambda_{\min}^2(L_1)} + 1 \quad (29)$$

$$c_2 > (c_1 - 1)\lambda_{\max}^2(L_1) + \lambda_{\max}(L_1)$$

then we have

$$\dot{V} < \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P - I_{n_\eta})] \tilde{\eta} + \gamma_1 e^{-\gamma_2 t} \quad (30)$$

From (15), it turns out that

$$\dot{V} < -\tilde{\eta}^T (I_n \otimes \alpha_1 P) \tilde{\eta} + \gamma_1 e^{-\gamma_2 t} \quad (31)$$

Inspired by Theorem 1 in [22], we define a new function $W = V + \frac{\gamma_1}{\gamma_2} e^{-\gamma_2 t}$ and calculate its derivative that $\dot{W} < 0$. Hence, the dual adaptive laws θ_i and ϕ_i are bounded. We calculate $\dot{V} + \alpha_1 V$ under (15) to have

$$\begin{aligned} \dot{V} + \alpha_1 V &< \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P - I_{n_\eta} + \alpha_1 P)] \tilde{\eta} \\ &+ \alpha_1 \varepsilon_1 + \gamma_1 e^{-\gamma_2 t} < \varepsilon_2 \end{aligned} \quad (32)$$

where $\varepsilon_1 = \sum_{i=1}^n \frac{(\theta_i - c_1)^2 + (\phi_i - c_2)^2}{2}$, and ε_1 is also bounded when θ_i and ϕ_i are bounded. $\varepsilon_2 = \alpha_1 \bar{\varepsilon}_1 + \gamma_1 > 0$, and $\bar{\varepsilon}_1$ represents the upper bound of ε_1 .

For $t \in D_a(0, \infty)$ under DoS attacks, choosing the same Lyapunov function candidate in (18) implies

$$\begin{aligned} \dot{V} - \alpha_2 V &= \tilde{\eta}^T [I_n \otimes (PA^\eta + (A^\eta)^T P - \alpha_2 P)] \tilde{\eta} \\ &- e_\eta^T (c_2 \Theta \otimes I_{n_\eta}) e_\eta - \alpha_2 \varepsilon_1 + \gamma_1 e^{-\gamma_2 t} \end{aligned} \quad (33)$$

Substitution (16) into (33)

$$\dot{V} - \alpha_2 V < \gamma_1 e^{-\gamma_2 t} < \varepsilon_2 \quad (34)$$

Inspired by Lemma 2 in [11], if (32) and (33) hold then

$$\begin{aligned} V(t) &\leq \varepsilon_2 \int_{t_0}^T e^{\alpha_2 |D_a(s,T)| - \alpha_1 |D_w(s,T)|} ds \\ &+ e^{-\alpha_1 |D_w(t_0,T)| + \alpha_2 |D_a(t_0,T)|} V(t_0) \end{aligned} \quad (35)$$

Assumption 5 leads to

$$\begin{aligned} |D_w(s,T)| &= T - s - |D_a(s,T)|, \\ &- \alpha_1 |D_w(s,T)| + \alpha_2 |D_a(s,T)| \\ &\leq -\rho(T - s) + (\alpha_1 + \alpha_2) \varepsilon_a \end{aligned} \quad (36)$$

where $\rho = \alpha_1 - \frac{\alpha_1 + \alpha_2}{R_d}$. If R_d satisfied (17), it obtains from (35) that

$$\begin{aligned} V(t) &\leq e^{(\alpha_1 + \alpha_2) \varepsilon_a} \left(\varepsilon_2 \int_{t_0}^T e^{-\rho(T-s)} ds + e^{-\rho(T-t_0)} V(t_0) \right) \\ &= e^{(\alpha_1 + \alpha_2) \varepsilon_a} \left(\frac{\varepsilon_2}{\rho} (1 - e^{-\rho T}) + e^{-\sqrt{\rho}(T-t_0)} V(t_0) \right). \end{aligned} \quad (37)$$

Since given parameters α_1, α_2 and matrix P in Theorem 1 and known parameters ϵ_a and R_d in Assumption 5 are bounded, the observer error $\tilde{\eta}_i$ is also bounded by

$$\limsup_{t \rightarrow \infty} \|\tilde{\eta}_i\|^2 \leq \frac{\epsilon_3}{\lambda_{\min}(P)\rho} \quad (38)$$

where $\epsilon_3 = e^{(\alpha_1 + \alpha_2)\epsilon_a}\epsilon_2 - \rho\underline{\epsilon}_1 > \epsilon_2 - \rho\underline{\epsilon}_1 > 0$, and $\underline{\epsilon}_1$ representing the initial value of ϵ_1 is also the lower bound of ϵ_1 .

To demonstrate the Zeno behavior can be avoided, taking the derivations of $\phi_i\theta_i e_{\eta_i}^T e_{\eta_i}$ and $\psi_i^T \psi_i + \gamma_1 e^{-\gamma_2 t}$

$$\begin{aligned} \frac{d(\phi_i\theta_i e_{\eta_i}^T e_{\eta_i})}{dt} &= \dot{\phi}_i\theta_i e_{\eta_i}^T e_{\eta_i} + \phi_i\dot{\theta}_i e_{\eta_i}^T e_{\eta_i} \\ &\quad + 2\phi_i\theta_i e_{\eta_i}^T (A^\eta e_{\eta_i} - \theta_i K_i^\eta \psi_i) \\ &\leq (\dot{\phi}_i + \dot{\theta}_i + 2\|A^\eta\| + 1)\phi_i\theta_i e_{\eta_i}^T e_{\eta_i} \\ &\quad + (\phi_i\theta_i^3 \|K_i^\eta\|^2)\psi_i^T \psi_i \\ &\leq (\dot{\phi}_i + \dot{\theta}_i + 2\|A^\eta\| + 1)\phi_i\theta_i e_{\eta_i}^T e_{\eta_i} \\ &\quad + (\phi_i\theta_i^3 \|K_i^\eta\|^2)(\psi_i^T \psi_i + \gamma_1 e^{-\gamma_2 t}) \end{aligned} \quad (39)$$

$$\frac{d(\psi_i^T \psi_i + \gamma_1 e^{-\gamma_2 t})}{dt} = 2A^\eta \psi_i^T \psi_i - \gamma_1 \gamma_2 e^{-\gamma_2 t} \quad (40)$$

where $\|\star\|$ denotes the norm of matrix \star . Defining function $J_i = \frac{\phi_i\theta_i e_{\eta_i}^T e_{\eta_i}}{J_i^f}$ where $J_i^f = \psi_i^T \psi_i + \gamma_1 e^{-\gamma_2 t} > 0$ in any finite time with or without DoS attacks. The derivation of J_i under Assumption 1 satisfies that

$$\begin{aligned} \dot{J}_i &\leq \frac{\frac{d(\phi_i\theta_i e_{\eta_i}^T e_{\eta_i})}{dt}(J_i^f) - (\phi_i\theta_i e_{\eta_i}^T e_{\eta_i})\frac{d(J_i^f)}{dt}}{J_i^f J_i^f} \\ &\leq (\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\|)J_i + (\bar{\phi}_i\bar{\theta}_i^3 \|K_i^\eta\|^2) \\ &\quad - \frac{(\phi_i\theta_i e_{\eta_i}^T e_{\eta_i})(2A^\eta \psi_i^T \psi_i - \gamma_1 \gamma_2 e^{-\gamma_2 t})}{J_i^f J_i^f} \\ &\leq (\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\| + \gamma_2)J_i + (\bar{\phi}_i\bar{\theta}_i^3 \|K_i^\eta\|^2) \end{aligned} \quad (41)$$

It follows from (41) that

$$J_i \leq \frac{(\bar{\phi}_i\bar{\theta}_i^3 \|K_i^\eta\|^2)}{(\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\| + \gamma_2)} (e^{(\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\| + \gamma_2)\tau} - 1) \quad (42)$$

where $\tau \leq t_i^{k+1} - t_i^k$ is the smallest time interval under condition of (9).

$$\tau = \frac{\ln\left(\frac{\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\| + \gamma_2}{\bar{\phi}_i\bar{\theta}_i^3 \|K_i^\eta\|^2} + 1\right)}{\dot{\phi}_i + \dot{\theta}_i + 1 + 2\|A^\eta\| + \gamma_2} > 0$$

Therefore, the Zeno behavior is avoided for any finite time, which completes the proof. \square

Remark 3. Compared with the resilient observers designed under DoS attack, the eigenvalues of the Laplacian matrix of global topology are required in [14], [15]. We remove this constraint to fully decouple the two layers and broaden the application scopes of the proposed distributed resilient observer. Note that the system matrices of the exosystem are removed in [11] by online learning when Assumption 1 is stricter such that the eigenvalues of matrix A_i^v lie on the imaginary axis. Based on our previous work in [23], the exosystem system matrices can also be removed in the resilient observer.

B. Optimal TVF Control based on RL

We define n estimated reference systems consisting of the observed exosystem and TVF (3) as

$$\dot{\hat{v}}_i = A_i^v \hat{v}_i, \quad \hat{y}_i^v = C_i^v \hat{v}_i \quad i = 1, 2, \dots, n \quad (43)$$

where $\hat{v}_i = \text{col}(\hat{\eta}_i, f_i) \in \mathbb{R}^{n^v}$ is estimated system state. $A_i^v = \text{blkdiag}(A^\eta, A_i^f)$ and $C_i^v = [C_i^\eta, C_i^f]$ are system state and output matrices.

In the physical system layer, the optimal formation output controller is designed as

$$\begin{aligned} u_i &= -K_i^x x_i - K_i^z z_i \\ \dot{z}_i &= \Sigma_{i1} z_i + \Sigma_{i2} \hat{e}_i \\ \hat{e}_i &= y_i - C^\eta \hat{\eta}_i - C_i^f f_i \end{aligned} \quad (44)$$

where z_i and $(\Sigma_{i1}, \Sigma_{i2})$ are defined in (6) and (7). The optimal control law $K_i = [K_i^x, K_i^z]$ is obtained by Algorithm 1. Inspired by the persistence of excitation (PE) condition [19] for online learning, a new rank condition is given in Condition 1 for Algorithm 1.

Algorithm 1 Online off-policy RL algorithm

1. Apply an admissible control policy $u_i^k = -K_i^0 \varsigma_i + e_{di}$ with a noise term e_{di} for agent i and collect online data until the rank condition (45) is meet.
 2. $k \leftarrow 0$
 3. Repeat
 4. $k \leftarrow k + 1$
 5. Evaluate policy by solve P_i^k, K_i^{k+1} from (57).
 6. Until $\|K_i^{k+1} - K_i^k\| < \epsilon$, where ϵ is a small constant.
 7. $K_i^* \leftarrow K_i^{k+1}$
 8. Use $u_i = -K_i^* \varsigma_i$ as learned control policy
-

Condition 1. *There is a sufficiently large integer $l > 0$ such that*

$$\text{rank}(\Xi_k) = \frac{n_i^\varsigma(n_i^\varsigma + 1)}{2} + m_i n_i^\varsigma + n_i^\varsigma n_i^v \quad (45)$$

where $\varsigma_i = \text{col}(x_i, z_i) \in \mathbb{R}^{n_i^\varsigma}$,

$$\begin{aligned} \Xi_k &= \begin{bmatrix} \Gamma_{\varsigma_i \varsigma_i}, -2I_{\varsigma_i \varsigma_i} \left(I \otimes \left(K_i^k \right)^T R_i \right) - 2I_{\varsigma_i u_i} (I \otimes R_i), \\ -2I_{\varsigma_i v_i} \end{bmatrix} \\ \Gamma_{\varsigma_i \varsigma_i} &= \begin{bmatrix} \varsigma_i \otimes \varsigma_i|_{t_1}^{t_1+\Delta T}, \varsigma_i \otimes \varsigma_i|_{t_2}^{t_2+\Delta T}, \dots, \\ \varsigma_i \otimes \varsigma_i|_{t_l}^{t_l+\Delta T} \end{bmatrix}^T \\ I_{\varsigma_i \varsigma_i} &= \begin{bmatrix} \int_{t_1}^{t_1+\Delta T} \varsigma_i \otimes \varsigma_i dt, \int_{t_2}^{t_2+\Delta T} \varsigma_i \otimes \varsigma_i dt, \dots, \\ \int_{t_l}^{t_l+\Delta T} \varsigma_i \otimes \varsigma_i dt \end{bmatrix}^T \\ I_{\varsigma_i u_i} &= \begin{bmatrix} \int_{t_1}^{t_1+\Delta T} \varsigma_i \otimes u_i dt, \int_{t_2}^{t_2+\Delta T} \varsigma_i \otimes u_i dt, \dots, \\ \int_{t_l}^{t_l+\Delta T} \varsigma_i \otimes u_i dt \end{bmatrix}^T \\ I_{\varsigma_i v_i} &= \begin{bmatrix} \int_{t_1}^{t_1+\Delta T} \varsigma_i \otimes \hat{v}_i dt, \int_{t_2}^{t_2+\Delta T} \varsigma_i \otimes \hat{v}_i dt, \dots, \\ \int_{t_l}^{t_l+\Delta T} \varsigma_i \otimes \hat{v}_i dt \end{bmatrix}^T \end{aligned} \quad (46)$$

for $0 \leq t_1 < t_2 < \dots < t_l$, where K_i^k is i th agent's update policy matrix of the k th step.

In Condition 1, $[t_1, t_l]$ indicates the minimum data collection time interval that satisfies the PE condition.

Now, we are ready to present our second result in this paper.

Theorem 2. *Consider MAS (1) with exosystem (2) and TVF system (3) under DoS attack satisfying Assumptions 1-5. The optimal resilient TVF control problem defined in (7) is solved by event-based observers (8) at cyber-layer and optimal formation controller (44) with RL Algorithm 1 at physical system layer.*

Proof. Substituting controller (44) into augmented system (4) implies

$$\begin{aligned} \dot{\varsigma}_i &= \tilde{A}_i^\varsigma \varsigma_i + \tilde{B}_i^\varsigma u_i + \tilde{B}_i^c \hat{v}_i + M_{i1} \\ &= \tilde{A}_i^c \varsigma_i + \tilde{B}_i^c v_i + M_{i2} \\ e_i &= \tilde{C}_i^c \varsigma_i - C_i^v v_i \end{aligned} \quad (47)$$

where

$$\begin{aligned}\tilde{A}_i^\zeta &= \begin{bmatrix} \tilde{A}_i & 0 \\ \Sigma_{i2}\tilde{C}_i & \Sigma_{i1} \end{bmatrix}, \quad \tilde{B}_i^\zeta = \begin{bmatrix} \tilde{B}_i \\ 0 \end{bmatrix}, \quad \tilde{B}_i^c = \begin{bmatrix} \tilde{F}_i \\ -\Sigma_{i2}C_i^v \end{bmatrix} \\ \tilde{A}_i^c &= \begin{bmatrix} \tilde{A}_i - \tilde{B}_iK_i^x & -\tilde{B}_iK_i^z \\ \Sigma_{i2}\tilde{C}_i & \Sigma_{i1} \end{bmatrix}, \quad M_{i1} = \begin{bmatrix} -\tilde{F}_i\tilde{v}_i \\ 0 \end{bmatrix} \\ M_{i2} &= \begin{bmatrix} 0 \\ -\Sigma_{i2}C_i^v\tilde{v}_i \end{bmatrix}, \quad \tilde{C}_i^c = [\tilde{C}_i \ 0]\end{aligned}$$

Based on Lemma 1.20 and Lemma 1.27 in [8], perturbed \tilde{A}_i^ζ is Hurwitz if A_i^c is Hurwitz under Assumption 4. There is a unique solution (Π_i^x, Π_i^z) to equation

$$\begin{aligned}\Pi_i^x A_i^v &= (\tilde{A}_i - \tilde{B}_i K_i^x) \Pi_i^x - \tilde{B}_i K_i^z \Pi_i^z + \tilde{F}_i \\ \Pi_i^z A_i^v &= \Sigma_{i1} \Pi_i^z + \Sigma_{i2} (\tilde{C}_i \Pi_i^x - C_i^v)\end{aligned}\tag{48}$$

The solution satisfies that

$$0 = \tilde{C}_i \Pi_i^x - C_i^v$$

Substituting a new variable $\Pi_i^v = \text{col}(\Pi_i^x, \Pi_i^z)$, into (48) yields

$$\Pi_i^v A_i^v = \tilde{A}_i^c \Pi_i^v + \tilde{B}_i^c \quad 0 = \tilde{C}_i^c \Pi_i^v - C_i^v\tag{49}$$

Equation (49) and output regulation equation (6) are equivalent with $U_i^v = -K_i^x \Pi_i^x - K_i^z \Pi_i^z$. Consider system (4) and (49), the dynamics of the state error $\tilde{\zeta}_i = \zeta_i - \Pi_i^v v_i$ and formation error e_i are

$$\dot{\tilde{\zeta}}_i = \tilde{A}_i^c \tilde{\zeta}_i + \tilde{B}_i^c v_i + M_{i2} - \Pi_i^v A_i^v v_i = \tilde{A}_i^c \tilde{\zeta}_i + M_{i2}\tag{50}$$

$$e_i = \tilde{C}_i^c \tilde{\zeta}_i - C_i^v v_i = \tilde{C}_i^c \tilde{\zeta}_i + (\tilde{C}_i^c \Pi_i^v - C_i^v) v_i = \tilde{C}_i^c \tilde{\zeta}_i\tag{51}$$

We rewrite the closed-loop formation error system as

$$\dot{\tilde{\zeta}}_i = \tilde{A}_i^c \tilde{\zeta}_i + M_{i2}, \quad e_i = \tilde{C}_i^c \tilde{\zeta}_i$$

Since M_{i1} , M_{i2} and \tilde{e}_i are bounded based on Theorem 1, the optimal resilient TVF control problem defined in (7) can be addressed in terms of error system

$$\dot{\tilde{\zeta}}_i = \tilde{A}_i^c \tilde{\zeta}_i + \tilde{B}_i^c \tilde{u}_i, \quad \tilde{u}_i = -K_i \tilde{\zeta}_i, \quad e_i = \tilde{C}_i^c \tilde{\zeta}_i\tag{52}$$

by solving the following RL policy iteration equation [24]

$$\begin{aligned}& \zeta_i^T(t + \Delta T) P_i^k \zeta_i(t + \Delta T) - \zeta_i^T(t) P_i^k \zeta_i(t) \\ &= \int_t^{t+\Delta T} \left[-\zeta_i^T \left(Q_i + (K_i^k)^T R_i K_i^k \right) \zeta_i \right. \\ & \quad \left. + 2(u_i - u_i^k)^T R_i K_i^{k+1} \zeta_i + 2\tilde{v}_i^T (\tilde{B}_i^c)^T P_i^k \zeta_i \right] dt\end{aligned}\tag{53}$$

where ΔT is the time interval. Driven by continuously collected data, model information can be learned, and control policies can be continuously optimized through iteration.

Note u_i in (54) and (44) are consistent.

$$\begin{aligned}
 u_i &= \tilde{u}_i + U_i^v v_i \\
 &= -K_i^x \tilde{x}_i - K_i^z \tilde{z}_i + U_i^v v_i \\
 &= -K_i^x x_i - K_i^z z_i \\
 &= -K_i \varsigma_i
 \end{aligned} \tag{54}$$

If the MAS model is known, the optimal control law is $K_i = R_i^{-1}(\tilde{B}_i^c)^T P_i$, where P_i is the solution of Riccati equation

$$(\tilde{A}_i^c)^T P_i + P_i \tilde{A}_i^c - P_i \tilde{B}_i^c R_i^{-1} (\tilde{B}_i^c)^T P_i + Q_i = 0 \tag{55}$$

for given $R_i = R_i^T > 0$ and $Q_i = Q_i^T > 0$ matrices. When the model information of MAS is unknown, that is, no knowledge of \tilde{A}_i^c and \tilde{B}_i^c is available, then policy iteration RL provides a solution for (55) based on the following equation from (53).

$$\begin{aligned}
 &\varsigma_i^T(t + \Delta T) P_i^k \varsigma_i(t + \Delta T) - \varsigma_i^T(t) P_i^k \varsigma_i(t) \\
 &= \int_t^{t+\Delta T} \left[-(\varsigma_i^T \otimes \varsigma_i^T) \text{vec} \left(Q_i + (K_i^k)^T R_i K_i^k \right) \right. \\
 &\quad + 2 \left((\varsigma_i^T \otimes \varsigma_i^T) \left(I \otimes (K_i^k)^T R_i \right) \right. \\
 &\quad + (\varsigma_i^T \otimes u_i^T) (I \otimes R_i) \left. \right) \text{vec} \left(K_i^{k+1} \right) \\
 &\quad \left. + 2 (\varsigma_i^T \otimes \hat{v}_i^T) \text{vec} \left((\tilde{B}_i^c)^T P_i^k \right) \right] dt
 \end{aligned} \tag{56}$$

where notation vec of matrix X represents $\text{vec}(X) = [X(1, :), X(2, :), \dots, X(n, :)]^T$.

From (56), we get

$$\Xi_k \begin{bmatrix} \text{vec} (P_i^k) \\ \text{vec} (K_i^{k+1}) \\ \text{vec} \left((\tilde{B}_i^c)^T P_i^k \right) \end{bmatrix} = \Upsilon_k \tag{57}$$

where $\Upsilon_k = -I_{\varsigma_i \varsigma_i} \text{vec} (Q_i + (K_i^k)^T R_i K_i^k)$, and Ξ_k is defined in (46).

Theorem 3 [24] and Theorem 2 [19] indicate that matrices P_i and K_i from Algorithm 1 converge to the optimal solutions P_i^* and K_i^* of the Riccati equation (55). The unique solution for K_i^{k+1} , P_i^k and $(\tilde{B}_i^c)^T P_i^k$ from (56) can be solved by the least squares method under rank condition (45). The formation output error is bounded, and it converges to zero when the observation error converges to zero. Therefore,

Algorithm 1 minimizes the performance index in (7) of the optimal resilient TVF control problem, which completes the proof. \square

Remark 4. Note that z_i is a dynamic compensator to improve system robustness against uncertainties and disturbances. Algorithm 1 is an online off-policy RL to solve the optimal resilient TVF control problem when the system model information is unknown. An initial admissible control policy is needed and a noise term e_{di} is introduced for stimulating deviation of equation (56) to meet the rank condition given in Condition 1. Compared to our previous work [23], the newly added rank condition can automatically adjust online data collection time, thereby improving online learning and optimization performance.

IV. EXPERIMENTAL VERIFICATION

Experiments of the area scanning formation of three unmanned ground vehicles (UGVs) and one virtual leader are conducted to verify the proposed strategies.

A virtual leader is designed with state $x_0 = \text{col}(x_c, y_c, \dot{x}_c, \dot{y}_c)$ and output $y_0 = \text{col}(x_c, y_c)$, where (x_c, y_c) is the position of virtual leader. The system state and output matrices are

$$A_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, C_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (58)$$

The wind disturbance is considered along the virtual leader's velocity direction with $A_w = \text{diag}(0, 0)$ which represents a constant main wind dynamics. In order to scan the area in an 'S' shape, a TVF system is designed for i th agent with state $f_i = [f_i^x, f_i^y, \dot{f}_i^x, \dot{f}_i^y, 1]^T$, where $\dot{f}_i^x = 0$ and $\dot{f}_i^y = s_i \sin(\omega_f t)$ describe a uniform linear motion along the X-direction and a sinusoidal motion along the Y-direction with an amplitude s_i and scanning rate ω_f . To model an 'S' shape scan task along the movement direction of the virtual leader, the TVF system matrices in (3) are given as

$$A_i^f = (I_2 \otimes D) A_w^f (I_2 \otimes D^T), C_i^f = \begin{bmatrix} 1 & 0 & 0 & 0 & c_i^x \\ 0 & 1 & 0 & 0 & c_i^y \end{bmatrix}$$

$$A_w^f = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -\omega_f^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, D = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \quad (59)$$

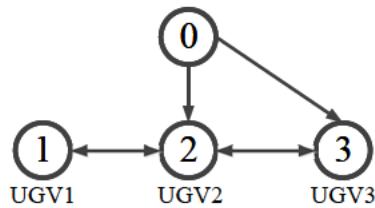


Fig. 2. Communication topology

where α is the angle between the virtual leader's velocity direction and the X-axis of the inertial system. Parameter (c_i^x, c_i^y) is the displacement deviation based on the scanning range.

The nonlinear model of UGVs is given as

$$\dot{p}_i^x = V_i \cos(\psi_i), \dot{p}_i^y = V_i \sin(\psi_i), \dot{V}_i = a_i, \dot{\psi}_i = r_i \quad (60)$$

where (p_i^x, p_i^y) is i th UGV's position. Variables V_i , a_i , ψ_i and r_i denote its velocity, acceleration, yaw angle and angular rate. Notice that the nonlinear UGV (60) is subject to the nonholonomic constraint. By introducing the following auxiliary variables (u_i^x, u_i^y)

$$\begin{bmatrix} a_i \\ r_i \end{bmatrix} = \begin{bmatrix} \cos(\psi_i) & \sin(\psi_i) \\ -\frac{1}{V_i} \sin(\psi_i) & \frac{1}{V_i} \cos(\psi_i) \end{bmatrix} \begin{bmatrix} u_i^x \\ u_i^y \end{bmatrix} \quad (61)$$

UGV system (60) is converted to

$$\dot{p}_i^x = v_i^x, \dot{p}_i^y = v_i^y, \dot{v}_i^x = u_i^x + d_i^x \omega_x, \dot{v}_i^y = u_i^y + d_i^y \omega_y \quad (62)$$

where, (v_i^x, v_i^y) represent the components of velocity V_i on the X and Y axes. $\omega = [\omega_x, \omega_y]$ denotes disturbance. The output defined in (1) is the position of i th UGV, that is $y_i = [p_i^x, p_i^y]^T$.

The parameters for experiments are taking as $x_0(t_0) = [0.2, 0, 0.01, 0]^T$ for virtual leader, $\omega_f = 0.1$, $s_{1,2,3} = 0.3$, $\alpha = 0$, $c_{1,2,3}^x = 0$, $c_1^y = 0$, $c_2^y = 0.7$, $c_3^y = -0.7$, $y_{1,2,3}^f(t_0) = [0, 0, 0.03, 0]^T$ for TVF. For disturbances, $\omega_0 = [0.1, 0]^T$, and the disturbance matrix is $d_{1,2,3}^x = (1, 1, 0.1)$ and $d_{1,2,3}^y(t_0) = (0.1, 0.1, 0.1)$. Only UGV2 and UGV3 are equipped with onboard cameras and all UGVs have wireless communication modules. An OptiTrack camera system is used for localization. Their communication topology is shown in Fig. 2, where the reference exosystem is indexed by the number 0, and agents 1-3 are real UGVs.

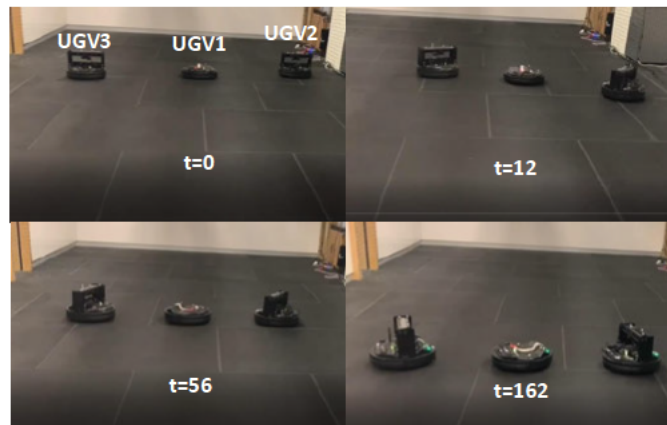


Fig. 3. Experiment on three-UGV cooperative area scanning formation under DoS attack

TABLE I
PERFORMANCE OF EVENT-TRIGGERED OBSERVER UNDER DOS

DoS parameter R_d	Average interval time	Frequency reduced
3.5	0.0884s	8 times
5	0.1077s	10 times
8	0.1345s	13 times
10	0.1779s	17 times

We use a typical DoS attack signal [13] as follows, which satisfies Assumption 5

$$\begin{aligned}
 d_c &= (c + 1) + 0.5c(c + 1) - \frac{(c + 1)}{R_d}, \\
 \sigma_c &= \frac{(c + 1)}{R_d}
 \end{aligned} \tag{63}$$

where $c = 1, 2, 3, \dots$, and $[d_c, d_c + \sigma_c)$ represents the c th DoS interval. The attack signal can be added to the communications module of UGVs by a real-time control software that deploys and validates real-time applications on hardware using Simulink. Note that we only consider DOS attacks on the observer communication channel used to exchange exosystem information, as shown in Fig. 1. The parameters of event-based observer (8) are set as $\alpha_1 = 0.1$, $\phi_i(t_0) = \theta_i(t_0) = 1.5$ and $\alpha_2 = 0.2$. Based on Theorems 1, the bound for R_d should meet $R_d > 3$. The photos of the experimental process are shown in Fig. 3 and the results of observation errors are given in Fig. 5 when $R_d = 3.5$. Table I also shows the results of resilient event-triggered strategy with different DoS attack parameter R_d . Figs. 6-10 demonstrate the results of formation control under $R_d = 3.5$, $R = 2I$ and $Q = 0.1I$ based on RL Algorithm 1.

The event-triggered interval of each agent is presented in Fig. 4 under DoS attack when $R_d = 3.5$.

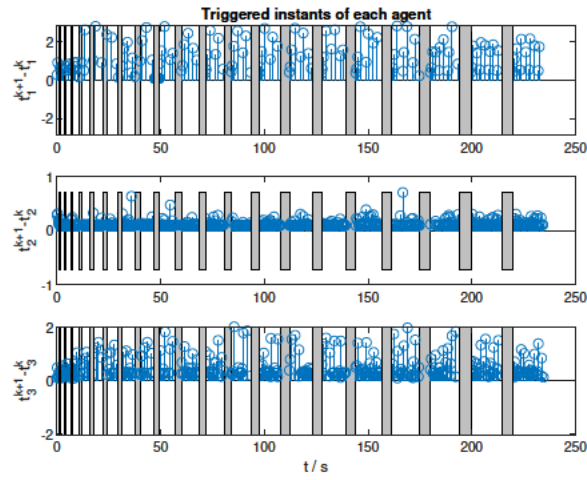


Fig. 4. Time instants of event-based observer for each agent under DoS attack

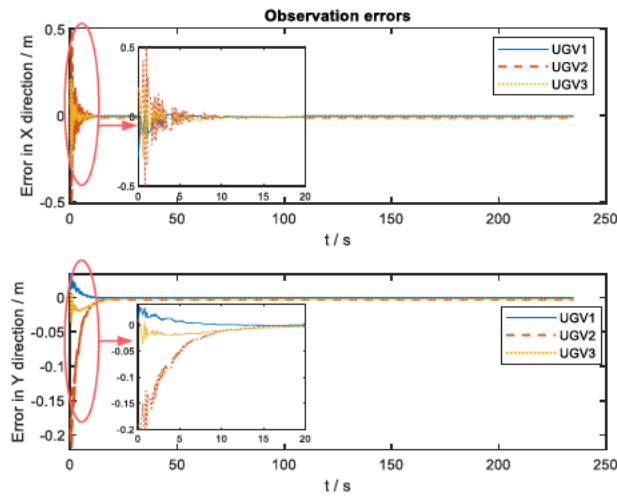


Fig. 5. Observation errors under DoS attack

The whole MAS has an average interval period of 0.0884 seconds. The error under the resilient observer converges to practically zero after 20 seconds, as shown in Fig. 5.

Several sets of comparative experiments were carried out when the DoS attack parameter R_d was 5, 8, and 10, and the event-triggered average interval times of the MAS were 0.1077 seconds, 0.1345 seconds, and 0.1779 seconds, respectively, as shown in Table I. The observation errors under different DoS attack parameters are all convergence to near zero within 20 seconds. The reduction in communication frequency is also given in Table I. Compared to the resilient observer design in [11] that cannot reduce the communication frequency, even in the worst case of the longest attack duration when $R_d = 3.5$ using

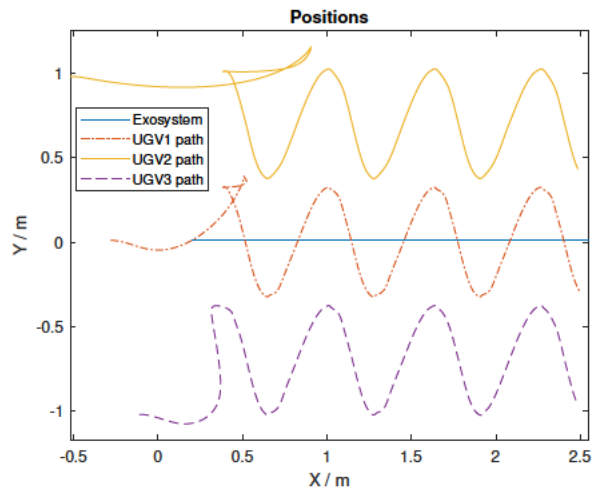


Fig. 6. Positions of agents under DoS attack

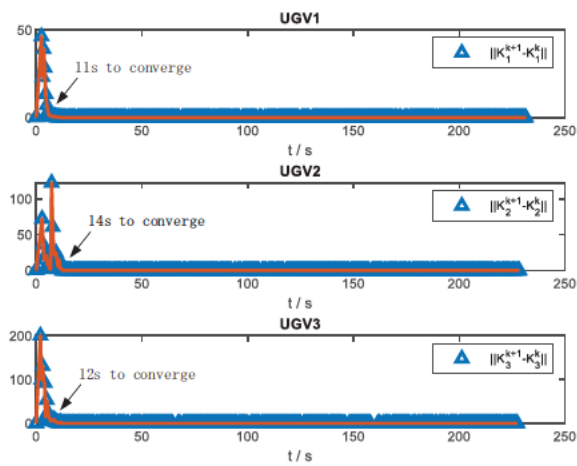


Fig. 7. Convergence trend of online off-policy RL Algorithm 1

the proposed observer, the frequency is still 8 times slower than the default communication frequency of the physical system. Unlike event-triggered controllers under attacks in [14], [15], our design is fully distributed and does not require global communication Laplacian matrix information. Therefore, the proposed observer can not only effectively increase communication security and save communication resources, but also remove the dependence on global communication information.

The UGV trajectories are shown in Fig. 3 and Fig. 6, where three UGVs track an exosystem and TVF

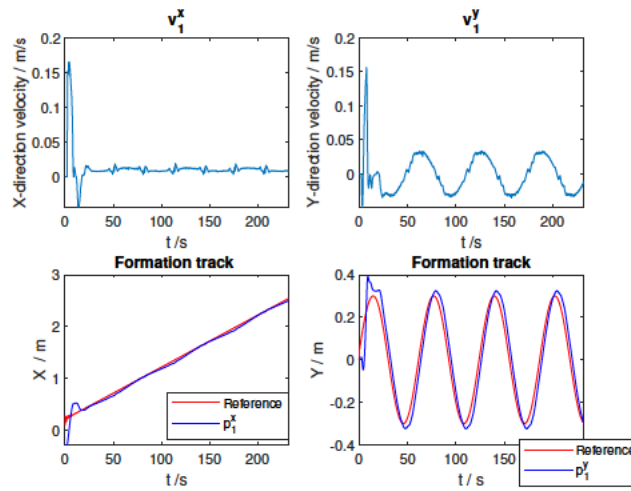


Fig. 8. TVF performance of UGV1

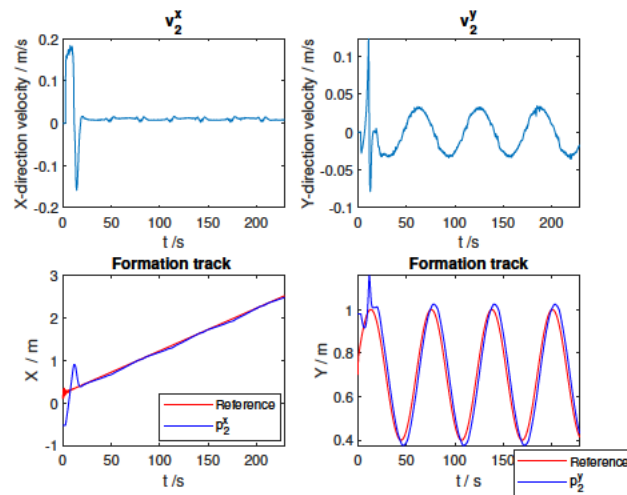


Fig. 9. TVF performance of UGV2

to move as an ‘S’ shape for area scanning. The online RL convergence trend of control law for each agent is given in Fig. 7. The norm errors $\|K_i^{k+1} - K_i^k\|$, $i = 1, 2, 3$ for UGVs converges to $\epsilon = 0.01$ at 11 seconds, 14 seconds and 12 seconds, respectively. The time-varying formation performance of three agents is displayed in Figs. 8-10. The positions of UGVs can practically track the designed reference formation system within 30 seconds under the DoS attack based on the resilient observer and online RL algorithm.

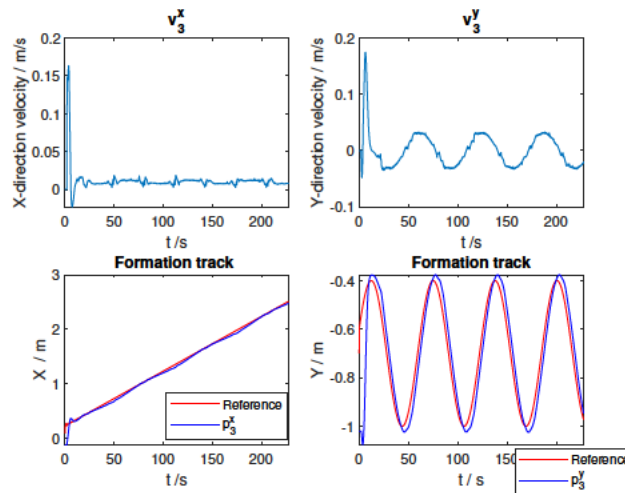


Fig. 10. TVF performance of UGV3

V. CONCLUSION

In this paper, a resilient two-layer time-varying formation control is provided for uncertain and unknown MAS under DoS attacks. For the cyber-layer, a fully distributed event-triggered observer is developed to remove global graph information. For the physical system layer, an optimal formation output control is proposed with off-policy RL to remove model information. Experiment results demonstrate the translation of our theoretical discoveries to applications such as collaborative area scanning. Our future works will consider various types of attacks such as false data injection attacks and zero-dynamics attacks, and develop resilient RL solutions to collision-free formation problems for MAS under attacks based on data-efficient RL and control barrier function [25].

REFERENCES

- [1] P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 161–175, 2021.
- [2] W. Xu, G. Hu, D. W. C. Ho, and Z. Feng, "Distributed secure cooperative control under denial-of-service attacks from multiple adversaries," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3458–3467, 2020.
- [3] Y. Fei, P. Shi, and C.-C. Lim, "Robust and collision-free formation control of multiagent systems with limited information," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi: 10.1109/TNNLS.2021.3112679.
- [4] Y.-S. Ma, W.-W. Che, C. Deng, and Z.-G. Wu, "Observer-based fully distributed containment control for MASs subject to DoS attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, doi: 10.1109/TSMC.2022.3189092.
- [5] Y. Sun, P. Shi, and C.-C. Lim, "Event-triggered adaptive leaderless consensus control for nonlinear multi-agent systems with unknown backlash-like hysteresis," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 15, pp. 7409–7424, 2021.

- [6] H. Cai and J. Huang, "Output based adaptive distributed output observer for leader–follower multiagent systems," *Automatica*, vol. 125, p. 109413, 2021.
- [7] X. Li, P. Shi, Y. Wang, and S. Wang, "Cooperative tracking control of heterogeneous mixed-order multiagent systems with higher-order nonlinear dynamics," *IEEE Transactions on Cybernetics*, vol. 52, no. 6, pp. 5498–5507, 2022.
- [8] J. Huang, *Nonlinear output regulation: theory and applications*, Philadelphia, PA: SIAM, 2004.
- [9] Y. Hua, X. Dong, L. Han, Q. Li, and Z. Ren, "Finite-time time-varying formation tracking for high-order multiagent systems with mismatched disturbances," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3795–3803, 2020.
- [10] B. Yan, P. Shi, C.-C. Lim, and C. Wu, "Robust formation control for multiagent systems based on adaptive observers," *IEEE Systems Journal*, pp. 1–12, 2021, doi:10.1109/JSYST.2021.3127579.
- [11] C. Deng, D. Zhang, and G. Feng, "Resilient practical cooperative output regulation for MASs with unknown switching exosystem dynamics under DoS attacks," *Automatica*, vol. 139, p. 110172, 2022.
- [12] R. Ma, P. Shi, and L. Wu, "Dissipativity-based sliding-mode control of cyber-physical systems under denial-of-service attacks," *IEEE Transactions on Cybernetics*, vol. 51, no. 5, pp. 2306–2318, 2021.
- [13] C. De Persis and P. Tesi, "Input-to-state stabilizing control under denial-of-service," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2930–2944, 2015.
- [14] Y. Xu, M. Fang, P. Shi, and Z.-G. Wu, "Event-based secure consensus of multiagent systems against DoS attacks," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3468–3476, 2020.
- [15] Y. Xu, M. Fang, Y.-J. Pan, K. Shi, and Z.-G. Wu, "Event-triggered output synchronization for nonhomogeneous agent systems with periodic denial-of-service attacks," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 6, pp. 1851–1865, 2021.
- [16] H. Liu, T. Ma, F. L. Lewis, and Y. Wan, "Robust formation control for multiple quadrotors with nonlinearities and disturbances," *IEEE Transactions on Cybernetics*, vol. 50, no. 4, pp. 1362–1371, 2018.
- [17] B. Pang, Z.-P. Jiang, and I. Mareels, "Reinforcement learning for adaptive optimal control of continuous-time linear periodic systems," *Automatica*, vol. 118, p. 109035, 2020.
- [18] Y. Gao, W. Wang, and N. Yu, "Consensus multi-agent reinforcement learning for Volt-VAR control in power distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3594–3604, 2021.
- [19] W. Gao, M. Mynuddin, D. C. Wunsch, and Z.-P. Jiang, "Reinforcement learning-based cooperative optimal output regulation via distributed adaptive internal model," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi:10.1109/TNNLS.2021.3069728.
- [20] Z. Li, M. Z. Chen, and Z. Ding, "Distributed adaptive controllers for cooperative output regulation of heterogeneous agents over directed graphs," *Automatica*, vol. 68, pp. 179–183, 2016.
- [21] Z. Qu, *Cooperative control of dynamical systems: applications to autonomous vehicles*, London, U.K.: Springer-Verlag, 2009.
- [22] Y.-Y. Qian, L. Liu, and G. Feng, "Distributed dynamic event-triggered control for cooperative output regulation of linear multiagent systems," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3023–3032, 2019.
- [23] B. Yan, P. Shi, C.-C. Lim, and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 5, pp. 2683–2704, 2022.
- [24] H. Modares, S. P. Nagesh Rao, G. A. D. Lopes, R. Babuška, and F. L. Lewis, "Optimal model-free output synchronization of heterogeneous systems using off-policy reinforcement learning," *Automatica*, vol. 71, pp. 334–341, 2016.

- [25] Y. Chen, A. Singletary, and A. D. Ames, “Guaranteed obstacle avoidance for multi-robot operations with limited actuation: a control barrier function approach,” *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 127–132, 2020.

Chapter 7

Conclusion

THIS thesis addresses a series of distributed formation control problems for heterogeneous multi-agent systems (MAS) under multiple constraints from interaction and physical systems. Applications for autonomous vehicles are also presented to verify the effectiveness of the proposed methods. This chapter summarizes the research presented in this thesis and introduces possible future work.

7.1 Summary

Formation control of MAS and its applications are considered in this thesis. The main research works can be summarized as follows:

- A robust L_2 bounded formation control strategy is proposed for uncertain heterogeneous MAS subject to switching topologies and multiple disturbances. Based on an adaptive observer and an internal model control, the distributed time-varying formation (TVF) tracking can be realized under switching topologies, and the influence of homogeneous disturbances and heterogeneous disturbances on the system performance can be compensated dynamically. The simulation results show that the scheme can improve the robustness of the system under multiple disturbances compared with the traditional control method.
- A solution of the nonlinear TVF control problem is provided for uncertain nonlinear heterogeneous MAS under limited communication bandwidth. To save network resources, a dual adaptive event-triggered observer is designed to estimate the reference exosystem, and the Zeno phenomenon can be excluded. Nonlinear formation control approaches are developed based on nonlinear output regulation control for MAS to perform collaborative tasks. The method is applied to a UAV-UGV MAS for simulations and a practical multi-UGV platform. Simulation and Experimental results demonstrate that the proposed control strategies are effective in providing a theoretical reference for the application of formation control in collaborative tracking and patrolling tasks in search and rescue operations.
- To achieve TVF for unknown MAS with an unknown exosystem in dynamic environments, a learning-based collision-free formation optimization strategy is proposed. Two reinforcement learning (RL)-based distributed observers are developed to learn exosystem dynamics and outputs. An off-policy RL algorithm and a collision-free optimization are proposed to minimize the non-quadratic formation objective function for unknown heterogeneous MAS. The results obtained from the simulation study verify the effectiveness and robustness of the proposed distributed observer and formation strategies.
- To deal with denial-of-service (DoS) attacks, a resilient two-layer formation control is provided for uncertain and unknown MAS. For the network layer, a fully

distributed event-triggered observer is developed to improve communication resilience without global graph information under DoS attacks. For the physical system layer, an optimal formation output control is proposed based on off-policy RL to remove model information with a new rank condition. Experimental results after applying to UGVs demonstrate our theoretical discoveries can be transformed into applications such as multi-vehicle collaborative area scanning.

7.2 Future works

There are a number of challenges remaining and deserving our future investigation:

1. **Develop security and safety-critical collaborative control frameworks under multiple attacks and failures**

Together with DoS attacks, other types of cyber-attacks, network-induced problems, and physical faults need to be considered, such as false data injection attacks [67], zero-dynamics attacks [68], network delay [69], packet loss [70], and sensor failures [71]. Our future works will focus on developing collaborative control and optimization frameworks to ensure communication security and formation safety for MAS with multiple cyber-attacks and faults.

2. **Enhance learning-based collaborative decision-making for heterogeneous systems**

In addition to multi-agent formation control, collaborative decision-making [72] is critical for heterogeneous MAS, even for human-machine heterogeneous systems. However, the unknown environments and unmodeled behaviors make the collaborative problem more complex and difficult to analyze. Combining deep RL with decision-making strategies makes it possible to explain unmodeled behaviors and achieve optimal decision-making in real-time [73]. Thus, deep RL-based decision-making strategies will be future studied to achieve deep and harmonious collaborations between heterogeneous entities.

3. **Shorten the gap between collaborative control theories and their real-world applications**

Currently, there is still a gap in the transformations from collaborative control theories into real-world applications, such as multi-vehicle autonomous driving

technologies. In the future, I will continue to pay attention to the collaborative control applications for multiple UAVs, UGVs, AUVs, and human-machine systems [74].

Bibliography

- [1] K. Zhang, Z. Yang, and T. Başar, “Multi-agent reinforcement learning: A selective overview of theories and algorithms,” *Handbook of Reinforcement Learning and Control*, pp. 321–384, 2021.
- [2] M. Wooldridge, *An introduction to multiagent systems*. Hoboken, NJ, USA: Wiley, 2009.
- [3] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, “Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications,” *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3826–3839, 2020.
- [4] G. Franzè, W. Lucia, and A. Venturino, “A distributed model predictive control strategy for constrained multi-vehicle systems moving in unknown environments,” *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 2, pp. 343–352, 2021.
- [5] K. Guo, X. Li, and L. Xie, “Ultra-wideband and odometry-based cooperative relative localization with application to multi-UAV formation control,” *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2590–2603, 2019.
- [6] M. Lu and L. Liu, “Leader-following consensus of multiple uncertain Euler-Lagrange systems with unknown dynamic leader,” *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4167–4173, 2019.
- [7] P. Shi and B. Yan, “A survey on intelligent control for multiagent systems,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 161–175, 2021.
- [8] H. Wang and M. Rubenstein, “Shape formation in homogeneous swarms using local task swapping,” *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 597–612, 2020.
- [9] C. Gao, Z. Wang, X. He, and D. Yue, “Sampled-data-based fault-tolerant consensus control for multi-agent systems: A data privacy preserving scheme,” *Automatica*, vol. 133, p. 109847, 2021.
- [10] H. Wang, W. Ren, W. Yu, and D. Zhang, “Fully distributed consensus control for a class of disturbed second-order multi-agent systems with directed networks,” *Automatica*, vol. 132, p. 109816, 2021.
- [11] Q. Li, J. Wei, Q. Gou, and Z. Niu, “Distributed adaptive fixed-time formation control for second-order multi-agent systems with collision avoidance,” *Information Sciences*, vol. 564, pp. 27–44, 2021.
- [12] N. Mathews, A. L. Christensen, R. O’Grady, and M. Dorigo, “Spatially targeted communication and self-assembly,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 2678–2679.
- [13] C. Chen, F. L. Lewis, K. Xie, S. Xie, and Y. Liu, “Off-policy learning for adaptive optimal output synchronization of heterogeneous multi-agent systems,” *Automatica*, vol. 119, p. 109081, 2020.
- [14] D. Wang, Z. Wang, Z. Wang, and W. Wang, “Design of hybrid event-triggered containment controllers for homogeneous and heterogeneous multiagent systems,” *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 4885–4896, 2020.

-
- [15] Y. Liu and Z. Wang, "Optimal output synchronization of heterogeneous multi-agent systems using measured input-output data," *Information Sciences*, vol. 582, pp. 462–479, 2022.
- [16] Z. Liu, D. Nojavanzadeh, A. Saberi, and A. A. Stoorvogel, "Scale-free collaborative protocol design for output synchronization of heterogeneous multi-agent systems with nonuniform communication delays," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 4, pp. 2882–2894, 2022.
- [17] D. Li, S. S. Ge, W. He, G. Ma, and L. Xie, "Multilayer formation control of multi-agent systems," *Automatica*, vol. 109, p. 108558, 2019.
- [18] Z. Han, K. Guo, L. Xie, and Z. Lin, "Integrated relative localization and leader–follower formation control," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 20–34, 2018.
- [19] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, 2015.
- [20] L. Li, P. Shi, R. K. Agarwal, C. K. Ahn, and W. Xing, "Event-triggered model predictive control for multiagent systems with communication constraints," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 5, pp. 3304–3316, 2019.
- [21] S.-L. Dai, S. He, X. Chen, and X. Jin, "Adaptive leader–follower formation control of nonholonomic mobile robots with prescribed transient and steady-state performance," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3662–3671, 2020.
- [22] Z. Han, K. Guo, L. Xie, and Z. Lin, "Integrated relative localization and leader–follower formation control," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 20–34, 2019.
- [23] X. Dong and G. Hu, "Time-varying formation tracking for linear multiagent systems with multiple leaders," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3658–3664, 2017.
- [24] B. Yan, P. Shi, C.-C. Lim, and C. Wu, "Robust formation control for multiagent systems based on adaptive observers," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3139–3150, 2022.
- [25] Y. Gong, G. Wen, Z. Peng, T. Huang, and Y. Chen, "Observer-based time-varying formation control of fractional-order multi-agent systems with general linear dynamics," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 1, pp. 82–86, 2019.
- [26] B. Yan, P. Shi, and C.-C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," *IEEE Transactions on Automation Science and Engineering*, 2021, doi: 10.1109/TASE.2021.3103877.
- [27] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader–follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, 2008.
- [28] Y. Huang and Z. Meng, "Bearing-based distributed formation control of multiple vertical take-off and landing UAVs," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 3, pp. 1281–1292, 2021.
- [29] B. Yan, P. Shi, C.-C. Lim, and Z. Shi, "Optimal robust formation control for heterogeneous multi-agent systems based on reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 5, pp. 2683–2704, 2022.
-

-
- [30] M. O'Neill, H. Yue, S. Nag, P. Grogan, and O. de Weck, "Comparing and optimizing the darpa system F6 program value-centric design methodologies," in *AIAA SPACE 2010 Conference & Exposition*, 2010, p. 8828.
- [31] T. Massey and Y. Shtessel, "Continuous traditional and high-order sliding modes for satellite formation control," *Journal of Guidance, Control, and Dynamics*, vol. 28, no. 4, pp. 826–831, 2005.
- [32] J. Lin, Z. Miao, H. Zhong, W. Peng, Y. Wang, and R. Fierro, "Adaptive image-based leader–follower formation control of mobile robots with visibility constraints," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 7, pp. 6010–6019, 2020.
- [33] J. Hu, P. Bhowmick, and A. Lanzon, "Group coordinated control of networked mobile robots with applications to object transportation," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8269–8274, 2021.
- [34] A. Vahidi and A. Sciarretta, "Energy saving potentials of connected and automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 822–843, 2018.
- [35] Y. Kang, D. Luo, B. Xin, J. Cheng, T. Yang, and S. Zhou, "Robust leaderless time-varying formation control for nonlinear unmanned aerial vehicle swarm system with communication delays," *IEEE Transactions on Cybernetics*, 2022, doi: 10.1109/TCYB.2022.3165007.
- [36] Y. Wu, Z. Zuo, Q. Han, Y. Wang, and H. Yang, "Formation control of wheeled mobile robots with multiple virtual leaders under communication failures," *IEEE Transactions on Control Systems Technology*, 2022, doi: 10.1109/TCST.2022.3175315.
- [37] L. Li, Y. Li, Y. Zhang, G. Xu, J. Zeng, and X. Feng, "Formation control of multiple autonomous underwater vehicles under communication delay, packet discreteness and dropout," *Journal of Marine Science and Engineering*, vol. 10, no. 7, p. 920, 2022.
- [38] H. Yu, P. Shi, and C. C. Lim, "Scalable formation control in stealth with limited sensing range," *International Journal of Robust and Nonlinear Control*, vol. 27, no. 3, pp. 410–433, 2017.
- [39] J. Jiang and Y. Jiang, "Leader-following consensus of linear time-varying multi-agent systems under fixed and switching topologies," *Automatica*, vol. 113, p. 108804, 2020.
- [40] J. Ma, X. Yu, L. Liu, H. Ji, and G. Feng, "Global cooperative output regulation of linear multiagent systems with limited bandwidth," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 2, pp. 1017–1028, 2022.
- [41] M. S. Mahmoud, M. M. Hamdan, and U. A. Baroudi, "Modeling and control of cyber-physical systems subject to cyber attacks: A survey of recent advances and challenges," *Neurocomputing*, vol. 338, pp. 101–115, 2019.
- [42] R. Alguliyev, Y. Imamverdiyev, and L. Sukhostat, "Cyber-physical systems and their security issues," *Computers in Industry*, vol. 100, pp. 212–223, 2018.
- [43] S. Jhaver, I. Birman, E. Gilbert, and A. Bruckman, "Human-machine collaboration for content regulation: The case of reddit automoderator," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 5, pp. 1–35, 2019.
-

-
- [44] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and industry 5.0—inception, conception and perception," *Journal of Manufacturing Systems*, vol. 61, pp. 530–535, 2021.
- [45] Q. Yang, J. Zhou, X. Chen, and J. Wen, "Distributed MPC-based secondary control for energy storage systems in a DC microgrid," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5633–5644, 2021.
- [46] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.
- [47] S. Xiong and Z. Hou, "Data-driven formation control for unknown MIMO nonlinear discrete-time multi-agent systems with sensor fault," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi: 10.1109/TNNLS.2021.3087481.
- [48] S. Zhang, L. Ma, and X. Yi, "Model-free adaptive control for nonlinear multi-agent systems with encoding-decoding mechanism," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 489–498, 2022.
- [49] W. Song, J. Feng, and S. Sun, "Data-based output tracking formation control for heterogeneous mimo multiagent systems under switching topologies," *Neurocomputing*, vol. 422, pp. 322–331, 2021.
- [50] B. Pang, Z.-P. Jiang, and I. Mareels, "Reinforcement learning for adaptive optimal control of continuous-time linear periodic systems," *Automatica*, vol. 118, p. 109035, 2020.
- [51] W. Zhao, H. Liu, F. L. Lewis, and X. Wang, "Data-driven optimal formation control for quadrotor team with unknown dynamics," *IEEE Transactions on Cybernetics*, vol. 52, no. 8, pp. 7889–7898, 2022.
- [52] 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.
- [53] W. Zhao, H. Liu, Y. Wan, and Z. Lin, "Data-driven formation control for multiple heterogeneous vehicles in air-ground coordination," *IEEE Transactions on Control of Network Systems*, 2022, doi: 10.1109/TCNS.2022.3181254.
- [54] W. Gao, M. Mynuddin, D. C. Wunsch, and Z.-P. Jiang, "Reinforcement learning-based cooperative optimal output regulation via distributed adaptive internal model," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi: 10.1109/TNNLS.2021.3069728.
- [55] Y. Gao, W. Wang, and N. Yu, "Consensus multi-agent reinforcement learning for volt-var control in power distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3594–3604, 2021.
- [56] M. Liu, Y. Wan, F. L. Lewis, and V. G. Lopez, "Adaptive optimal control for stochastic multiplayer differential games using on-policy and off-policy reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 12, pp. 5522–5533, 2020.
- [57] Y. Li, W. Gao, S. Huang, R. Wang, W. Yan, V. Gevorgian, and D. W. Gao, "Data-driven optimal control strategy for virtual synchronous generator via deep reinforcement learning approach," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 4, pp. 919–929, 2021.
-

-
- [58] S. B. Prathiba, G. Raja, K. Dev, N. Kumar, and M. Guizani, "A hybrid deep reinforcement learning for autonomous vehicles smart-platooning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 13 340–13 350, 2021.
- [59] Z. Yan and Y. Xu, "A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4599–4608, 2020.
- [60] T. T. Mac, C. Copot, D. T. Tran, and R. De Keyser, "Heuristic approaches in robot path planning: A survey," *Robotics and Autonomous Systems*, vol. 86, pp. 13–28, 2016.
- [61] H. Li, W. Liu, C. Yang, W. Wang, T. Qie, and C. Xiang, "An optimization-based path planning approach for autonomous vehicles using the DynEFA-artificial potential field," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 2, pp. 263–272, 2022.
- [62] Z. Pan, C. Zhang, Y. Xia, H. Xiong, and X. Shao, "An improved artificial potential field method for path planning and formation control of the multi-UAV systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 3, pp. 1129–1133, 2022.
- [63] Y. Fei, P. Shi, and C.-C. Lim, "Robust and collision-free formation control of multiagent systems with limited information," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, doi: 10.1109/TNNLS.2021.3112679.
- [64] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2016.
- [65] A. Singletary, K. Klingebiel, J. Bourne, A. Browning, P. Tokumar, and A. Ames, "Comparative analysis of control barrier functions and artificial potential fields for obstacle avoidance," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 8129–8136.
- [66] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 3420–3431.
- [67] A. Sargolzaei, B. C. Allen, C. D. Crane, and W. E. Dixon, "Lyapunov-based control of a nonlinear multiagent system with a time-varying input delay under false-data-injection attacks," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2693–2703, 2021.
- [68] G. Park, C. Lee, H. Shim, Y. Eun, and K. H. Johansson, "Stealthy adversaries against uncertain cyber-physical systems: Threat of robust zero-dynamics attack," *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 4907–4919, 2019.
- [69] N. Lin and Q. Ling, "An event-triggered consensus protocol for quantized second-order multi-agent systems with network delay and process noise," *ISA Transactions*, vol. 125, pp. 31–41, 2022.
- [70] L. Xu, Y. Mo, and L. Xie, "Distributed consensus over Markovian packet loss channels," *IEEE Transactions on Automatic Control*, vol. 65, no. 1, pp. 279–286, 2019.
-

-
- [71] Y. Yu, J. Guo, and Z. Xiang, "Distributed fuzzy consensus control of uncertain nonlinear multiagent systems with actuator and sensor failures," *IEEE Systems Journal*, 2021, doi: 10.1109/JSYST.2021.3098930.
- [72] P. Landgren, V. Srivastava, and N. E. Leonard, "Distributed cooperative decision making in multi-agent multi-armed bandits," *Automatica*, vol. 125, p. 109445, 2021.
- [73] L. Chen, Y. He, Q. Wang, W. Pan, and Z. Ming, "Joint optimization of sensing, decision-making and motion-controlling for autonomous vehicles: A deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 4642–4654, 2022.
- [74] I. Seeber, E. Bittner, R. O. Briggs, T. De Vreede, G.-J. De Vreede, A. Elkins, R. Maier, A. B. Merz, S. Oeste-Reiß, N. Randrup *et al.*, "Machines as teammates: A research agenda on ai in team collaboration," *Information & management*, vol. 57, no. 2, p. 103174, 2020.