Article

Intelligent Distributed Swarm Control for Large-Scale Multi-UAV Systems: A Hierarchical Learning Approach

Shawon Dey [ID] **and Hao Xu ***

Department of Electrical and Biomedical Engineering, University of Nevada, Reno, NV 89557, USA
* Correspondence: haoxu@unr.edu

Abstract: In this paper, a distributed swarm control problem is studied for large-scale multi-agent systems (LS-MASs). Different than classical multi-agent systems, an LS-MAS brings new challenges to control design due to its large number of agents. It might be more difficult for developing the appropriate control to achieve complicated missions such as collective swarming. To address these challenges, a novel mixed game theory is developed with a hierarchical learning algorithm. In the mixed game, the LS-MAS is represented as a multi-group, large-scale leader–follower system. Then, a cooperative game is used to formulate the distributed swarm control for multi-group leaders, and a Stackelberg game is utilized to couple the leaders and their large-scale followers effectively. Using the interaction between leaders and followers, the mean field game is used to continue the collective swarm behavior from leaders to followers smoothly without raising the computational complexity or communication traffic. Moreover, a hierarchical learning algorithm is designed to learn the intelligent optimal distributed swarm control for multi-group leader–follower systems. Specifically, a multi-agent actor–critic algorithm is developed for obtaining the distributed optimal swarm control for multi-group leaders first. Furthermore, an actor–critic–mass method is designed to find the decentralized swarm control for large-scale followers. Eventually, a series of numerical simulations and a Lyapunov stability proof of the closed-loop system are conducted to demonstrate the performance of the developed scheme.

Keywords: game theory; reinforcement learning; adaptive dynamic programming; LS-MAS

1. Introduction

The concept of swarming in multi-agent systems (MASs) has been adopted from the biological swarming behavior in nature ranging from bacteria to more advanced mammals [1]. Examples of swarming behavior include flock of birds [2], school of fish [3], and cooperation of ants [4], which form groups of MASs to achieve certain tasks such as threatening predators, foraging food, and energy-efficient flying during migration. Over the past few decades, the biological swarm system has been widely adopted by researchers [5–13]. A survey of the recent development in the control and optimization of swarm systems was presented in [14]. Aside from that, the cooperative control problem for the swarming system was studied in [15]. Skobelev et al. [16] proposed a prototype system using a swarm of unmanned aerial vehicles (UAVs), which includes coordinated flight plans with reconfiguration in the presence of disruptive events. In [17], a new formation control scheme with smooth distributed consensus control for multi-UAV systems was developed. The authors in [18] proposed intelligent flight decision techniques for UAVs for cooperative target tracking using an end-to-end, cooperative, multi-agent reinforcement learning scheme. Zhao et al. [19] developed a flocking obstacle avoidance algorithm with shared obstacle information. Despite all of these recent efforts, the traditional cooperative control of large-scale multi-agent systems (LS-MASs) suffers from the well-known "Curse of Dimensionality" [20] and communication reliability [21] problem.

To address these challenges, a mixed game theory is proposed and utilized to formulate the optimal swarming problem for large scale multi-agent systems. The overall structure of the developed mixed game theoretical swarming is shown in Figure 1. First, the LS-MAS is reformulated as a multi-group large-scale leader–follower system by dividing a large number of agents into several subgroups with a leader and a large number of followers in each group. The objective of each group is to achieve optimal distributed swarming behavior. To achieve desired swarming, a mixed game theory was developed by seamlessly integrating (1) the cooperative game [22,23] to ensure the collective swarming behavior among multi-group leaders, where the leaders from each group cooperate with other leaders to maintain the overall multi-group swarming while avoiding inter-group collision, (2) the Stackelberg game [24,25] to connect the leader with its corresponding followers, where a set of coupling functions is introduced by the Stackelberg game in order to maintain leader–follower cohesion in each group, which guarantees that the followers from each group successfully achieve the swarming behavior of their respective leaders, and (3) the mean field game (MFG) [26–28] between non-cooperative followers used to continue the collective swarming behavior from the leaders to large-scale followers. The MFG deal involves the "Curse of Dimensionality" and the communication complexity problem of a traditional cooperative game. A probability density function (PDF) (i.e., mass) is employed in the MFG to replace the large number of agents' state information. Since the mass function has the same dimension as the state space, it eliminates the correlation between rising computing complexity and the agents' increasing population by reducing the dimension of the cost function.
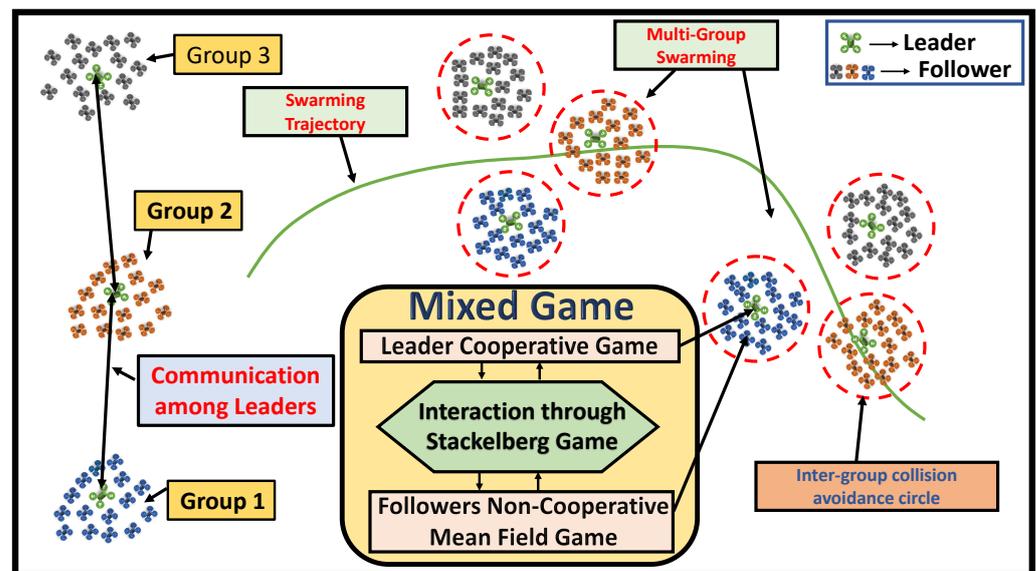

Figure 1. An illustration of mixed game-based leader–follower multi-group swarming.

However, to obtain the mass function, a new partial differential equation (PDE) (i.e., the Fokker–Planck–Kolmogorov (FPK) [29] equation), is adopted. Then, the optimal decentralized swarming control for large-scale followers in each group can be obtained by solving coupled Hamiltonian–Jacobi–Bellman (HJB) and FPK equations [30]. Meanwhile, the distributed optimal swarming control for multi-group leaders can be solved by using the multi-player cooperative HJB equation. However, it is very difficult and even impossible to solve those PDEs in real time due to its complexity and coupling. Hence, adaptive dynamic programming [31] and a reinforcement learning [32] algorithm were adopted. Specifically, a hierarchical learning structure was developed to obtain the optimal swarming control for the multi-group leader–follower system. This includes (1) multi-agent actor–critic-based distributed optimal swarm control for multi-group leaders and (2) actor–critic-mass (ACM)-based decentralized swarm control for large-scale followers. In ACM

learning-based control, the actor neural network (NN) is used to approximate the optimal decentralized control, the critic NN is used for approximating the optimal evaluation function, and the mass NN approximates the FPK solution. The main contributions of the article are as follows:

- A novel mixed game theory is developed with cooperative leaders and non-cooperative followers in order to achieve multi-group optimal swarming control which addresses the challenge of the curse of dimensionality and unrealistic communication.
- A hierarchical learning structure with actor–critic-based, leader-distributed swarming and actor–critic–mass-based, large-scale followers decentralized swarming is implemented in real time to learn the solution of the overall intelligent optimal swarming control.

The structure of this paper is as follows. Section 2 provides the significance of the developed algorithm, and Section 3 includes the problem formulation. In Section 4, the novel mixed game hierarchical learning-based intelligent optimal swarming control scheme is developed. Then, the numerical simulation is shown in Section 5 to demonstrate the effectiveness of the proposed design.

2. Significance of Mixed Game Theory-Based Intelligent Distributed Swarm Control

The traditional cooperative control [33–35] of multi-agent systems requires communication among all agents to achieve optimal control, which encounters issues with significant computational complexity and the requirement of low-latency communication in real time. In particular, when a swarm contains a massive population which is known as an LS-MAS, it suffers from the following challenges: (1) a high-quality communication network is required for exchanging information in an LS-MAS to achieve the conventional cooperative swarming behavior, although in practice, maintaining these communication networks is unrealistic, and (2) each agent must be aware of the states of the other agents to accomplish the desired multi-group swarming behavior. As the number of agents increases, the computational complexity problem increases substantially, which brings about the well-known "Curse of Dimensionality" [20] problem. Aside from that, most of the existing studies focus on the swarming behavior of single-group MASs with a limited number of agents while avoiding multi-group large-scale MASs, whose examples in a practical environment are ubiquitous. For instance, dividing an MAS into multiple groups can help it effectively handle multi-task missions, whereas a single group cannot, such as in predator formation during multi-prey hunting or cooperative searching for multiple objectives. Recently, several studies in MASs addressed this issue to some extent. Zhang et al. [36] developed a type of multi-group formation tracking control, where the agents are divided into several subgroups to form different desired sub-formations. A distributed impulsive control method with and without the input delay for multi-group formation tracking control was presented in [37]. Moreover, collision avoidance among multi-group UAVs was studied in [38]. In this paper, the developed mixed game theory-based distributed control addressed the challenges of the existing single-group traditional multi-agent cooperative control. At first, the large-scale multi-agent system is divided into multiple groups. Then, each group is reformulated with a leader and a large number of followers. The leader from each group plays the cooperative game to achieve the optimal swarming behavior for all the agents and ensure collision avoidance between the group of agents. The leader just needs to know the PDF information of the followers' group instead of knowing the state information of each follower. This reduces the dimension of the cost function significantly. However, the leaders still need to exchange the state information with other group leaders. The communication cost of exchanging information is still low compared with a large number of agents, as the agents are divided into a few of subgroups. On the other hand, this upper-level communication between leaders significantly ensures the optimal swarming behavior and collision avoidance of the group of agents. In addition, the followers from each group are guided by their respective leaders and play a non-cooperative game inside the group to reduce the computational complexity and communication problem significantly. To validate the effectiveness of the developed algorithm, the performance of the developed

method is further compared against the existing traditional cooperative control method in the simulation (Section 5.2).

3. Problem Formulation

Consider an LS-MAS being reformulated as a multi-group large-scale leader–follower system, with each group having a leader with a large number of followers. Then, assume there are M groups with M leaders as well as N_F^i being the number of followers in the ith group. Next, the dynamics of the leader and follower q in group i are defined as follows.
Leader:

$$dx_L^i = [F_a(x_L^i) + G_a(x_L^i)u_L^i]dt + \sqrt{2v^i}d\omega_L^i \tag{1}$$

where $x_L^i \in \mathbb{R}^m$ denotes the system state and $u_L^i \in \mathbb{R}^n$ denotes the control input. Moreover, $\omega_L^i \in \mathbb{R}^m$ denotes independent wiener processes which represents the environmental noise, and v^i is a non-negative parameter. The functions $F_a(.)$ and $G_a(.)$ represent the intrinsic dynamics of the leaders.
Follower:

$$dx_{F,q}^i = [F_s(x_{F,q}^i) + G_s(x_{F,q}^i)u_{F,q}^i]dt + \sqrt{2v^i}\omega_{F,q}^i \tag{2}$$

where $x_{F,q}^i \in \mathbb{R}^m$ is the state and $u_{F,q}^i \in \mathbb{R}^n$ is the control input of the follower q. Moreover, $\omega_{F,q}^i \in \mathbb{R}^m$ denotes the wiener process. The functions $F_s(.)$ and $G_s(.)$ represent the intrinsic dynamics of the followers.

3.1. Multi-Group Optimal Swarming Control Formulation

In this section, the mixed game theory is designed to formulate the optimal swarming control for multi-group large-scale leader–follower systems. Next, the details of the developed scheme are given as follows:

Collective Swarming among Multi-Group Leaders: To achieve the collective swarming behavior for all groups in the system, a predefined reference trajectory $x_d(t) \in \mathbb{R}^m$ is given to all leaders ahead of the mission. Next, the desired formation vector with respect to the reference trajectory for a group's ith leader can be denoted as $\lambda_L^i \in \mathbb{R}^m$. Then, the desired trajectory for the leader in group i is defined as follows:

$$x_{L,d}^i(t) = x_d(t) + \lambda_L^i \tag{3}$$

Then, the tracking error of the leader i is defined as $e_L^i(t) = x_L^i(t) - x_{L,d}^i(t)$ with the tracking error dynamic as follows:

$$\begin{aligned} de_L^i &= dx_L^i - dx_{L,d}^i \\ &= [F_{ar}(e_L^i) + G_{ar}(e_L^i)u_L^i]dt + \sqrt{2v^i}d\omega_L^i \end{aligned} \tag{4}$$

where $F_{ar}(e_L^i) = F_a(e_L^i + x_{L,d}^i) - (dx_{L,d}^i/dt)$ and $G_{ar}(e_L^i) = G_a(e_L^i + x_{L,d}^i)$. Additionally, the tracking error function for the leader in group i is given as

$$\Phi_{\text{SE,L}}(x_L^i) = \|x_L^i - x_{L,d}^i\|_{Q_{\text{SE,L}}}^2 \tag{5}$$

where $Q_{\text{SE,L}}$ is a matrix. To achieve the common goal for all groups (i.e., the collective swarming behavior), each leader needs to communicate with the leaders in the neighborhood groups to avoid inter-group collisions. Let $\mathcal{G} = \{\mathcal{A}, \mathcal{V}, \mathcal{E}\}$ is a graph that describes the connection among the leaders of M groups, with $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{M \times M}$ is an adjacency matrix (i.e., $\mathcal{A}^T = \mathcal{A}$). In addition, $\mathcal{V}(\mathcal{G}) = \{1, 2, ..., M\}$ denotes the set of leader vertices, and $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ is the set of edges. The element a_{ij} of the adjacency matrix is defined as $a_{ij} = a_{ji} = 1$ if $(i, j) \in \mathcal{E}$ (i.e., the leaders of group i and j are connected); otherwise, $a_{ij} = a_{ji} = 0$. Moreover, we assume that $a_{ii} = 0$. To define the neighborhood of each leader,

sensing and communicating distance is needed. Let $h > 0$ denote the communicating distance. Then, the neighborhood set $\mathcal{M}_L^i$ of the leader of group i is defined as follows:

$$\mathcal{M}_L^i(t) = \{j \in \mathcal{V} : \|x_L^i(t) - x_L^j(t)\|_{Q_1}^2 < h, j \neq i\} \tag{6}$$

where Q_1 is a positive definite matrix. Additionally, the communication between group leader i and j can be possible if $j \in \mathcal{M}_L^i(t)$. Furthermore, collision avoidance between leaders and their respective followers needs to be addressed for multi-group leader–follower swarming control in order to ensure safe path planning and guide the group to their desired swarming movement. The cost function for collision among leaders is

$$\Phi_{\mathrm{CA,L}}(x_L^i, x_L^{-i}) = w_{\mathrm{CA,L}} \sum_{j \in \mathcal{M}_L^i} a_{i,j} \left[\exp\{\|x_L^i - x_L^j\|_{Q_{\mathrm{CA,L}}}^2 - d_L^{i,j}\} - 1 \right]^{-1} \tag{7}$$

where $Q_{\mathrm{CA,L}}$ is a weighting matrix, $w_{\mathrm{CA,L}}$ is a weighting parameter, and $x_L^{-i} = \{x_L^j\}_{j \in \mathcal{M}_L^i}$. In addition, $d_L^{i,j}$ is the separating distance between leaders i and j, which is chosen so that it always ensures the avoidance of collision between two groups (i.e., collisions between any two leaders and their corresponding followers). Moreover, to achieve collective swarming behavior while avoiding group collision, the cohesion between the leader and large-scale followers needs to be maintained. In this regard, the leader–follower coupling function is introduced as follows.

Leader–Follower Coupling Functions for Swarming: To achieve a group swarming behavior, a set of coupling functions is defined by the Stackelberg game ([24]). The coupling function that forces the leader to keep cohesion with their corresponding followers is defined as follows:

$$\Phi_{\mathrm{CP,L}}(x_L^i, \rho^i) = w_{\mathrm{CP,L}} \|x_L^i - \mathbb{E}\{x_F^i\}\|_{Q_{\mathrm{CP,L}}}^2 \tag{8}$$

where $\rho^i(x_F^i, t)$ is the probability density function (PDF) of the large-scale followers' state in the ith group, $\mathbb{E}\{x_F^i\}$ is the expected value, and $w_{\mathrm{CP,L}}$ and $Q_{\mathrm{CP,L}}$ are the weighting matrix and weighting parameter, respectively. Now, the leader–follower swarming coupling function is

$$\Phi_{\mathrm{CP,F}}(x_{F,q}^i, x_L^i) = w_{\mathrm{CP,F}} \left[\exp\{r^i - \|x_{F,q}^i - x_L^i\|_{Q_{\mathrm{CP,F}}}^2\} - 1 \right]^{-1} \tag{9}$$

where $Q_{\mathrm{CP,F}}$ is a weighting matrix, $w_{\mathrm{CP,F}}$ is a weighting parameter, and r^i is the maximum safe distance of the ith group's followers. This distance r^i ensures inter-group collision avoidance by keeping the followers within the safe distance limit.

Leader–Follower Coupling Functions for Swarming: The followers from each group track their respective leaders in order to achieve the group swarming behavior. The tracking error of the follower q in group i is derived as $e_{F,q}^i(t) = x_{F,q}^i(t) - x_L^i(t)$, with the tracking error dynamics

$$de_{F,q}^i = \left[F_{sr}(e_{F,q}^i) + G_{sr}(e_{F,q}^i)u_{F,q}^i \right]dt + \sqrt{2\nu^i}d\omega_{F,q}^i \tag{10}$$

where

$$F_{sr}(e_{F,q}^i) = F_s(e_{F,q}^i + x_L^i) - F_a(x_L^i) \ ; \ G_{sr}(e_{F,q}^i) = G_s(e_{F,q}^i + x_L^i) - G_a(x_L^i)\frac{u_L^i}{u_{F,q}^i}$$

Now, the followers' tracking error cost function can be derived as follows:

$$\Phi_{\mathrm{SE,F}}(x_{F,q}^i) = \|x_{F,q}^i - x_L^i\|_{Q_{\mathrm{SE,F}}}^2 \tag{11}$$

where $Q_{SE,F}$ is a positive definite matrix. Then, the same group large scale follower collision avoidance function is as follows:

$$\Phi_{CA,F}(x^i_{F,q}, \rho^i) = w_{CA,F} \int_{x^i_F} \frac{\rho^i(x^i_F, t)}{[\varepsilon^2 + \|x^i_{F,q} - \mathbb{E}\{x^i_F\}\|^2_{Q_{CA,F}}]^\beta} dx^i_F \tag{12}$$

where $Q_{CA,F}$ and $w_{CA,F}$ are the weighting matrix and weighting parameter, respectively. Moreover, ε and β are positive constants. Furthermore, the cohesion function of the followers with their respective group center can be derived as follows:

$$\Phi_{C,F}(x^i_{F,q}, \rho^i) = w_{C,F}\left\{ \|x^i_{F,q} - \mathbb{E}\{x^i_F\}\|^2_{Q_{C,F}} \right\} \tag{13}$$

with the weighting matrix $Q_{C,F}$ and weighting parameter $w_{C,F}$. This function helps each follower from the same group to stay close to the other members of the group.

Overall Optimal Swarming Control for Multi-Group LS-MAS: The goal of the leaders is to achieve overall swarming control by minimizing the following cost function.

Leader:

$$V_L(x_L, \rho) = \sum_{i=1}^{M} V^i_L(x^i_L, x^{-i}_L, \rho^i) = \sum_{i=1}^{M} \mathbb{E}\left\{ \int_0^\infty \left[\Phi_L(x^i_L, x^{-i}_L, u^i_L) + \Phi_{CP,L}(x^i_L, \rho^i) \right] dt \right\} \tag{14}$$

where $\Phi_L(x^i_L, x^{-i}_L, u^i_L) = \Phi_{SE,L}(x^i_L) + \Phi_{CA,L}(x^i_L, x^{-i}_L) + \|u^i_L\|^2_{R^i_L}$, V^i_L is the cost function of the ith group leader and R^i_L is the positive weight matrix. The functions $\Phi_{SE,L}(x^i_L)$, $\Phi_{CA,L}(x^i_L, x^{-i}_L)$, and $\Phi_{CP,L}(x^i_L, \rho^i)$ derived in Equations (5), (7), and (8) are the tracking error, collision avoidance, and coupling function of the leader, respectively. Moreover, the objective of the follower q in the group i is to minimize the following cost function to achieve the group swarming behavior.

Follower:

$$V^i_{F,q}(x^i_{F,q}, x^i_L, \rho^i) = \mathbb{E}\left\{ \int_0^\infty \left[\Phi_F(x^i_{F,q}, x^i_L, u^i_{F,q}) + \Phi_{CA,F}(x^i_{F,q}, \rho^i) + \Phi_{C,F}(x^i_{F,q}, \rho^i) \right] dt \right\}$$
$$\text{s.t. } \|x^{i,o}_{F,q} - x^i_L\| < r^i \tag{15}$$

where $\Phi_F(x^i_{F,q}, x^i_L, u^i_{F,q}) = \Phi_{SE,F}(x^i_{F,q}) + \Phi_{CP,F}(x^i_{F,q}, x^i_L) + \|u^i_{F,q}\|^2_{R^i_{F,q}}$, $R^i_{F,q}$ is the positive weight matrix and $x^{i,o}_{F,q}$ is the initial state of the follower q. The functions $\Phi_{SE,F}(x^i_{F,q})$, $\Phi_{CP,F}(x^i_{F,q}, x^i_L)$, $\Phi_{C,F}(x^i_{F,q}, \rho^i)$, and $\Phi_{CA,F}(x^i_{F,q}, \rho^i)$ derived in Equations (9) and (11)–(13) are the follower tracking error, coupling, cohesion, and collision avoidance functions, respectively. Here, the initial states of the followers from all groups are subject to a constraint in order to keep the followers inside a specific region defined by the distance r^i. This distance is the radius of an enclosing circle which acts as a collision-avoiding region for the respective groups.

3.2. Mixed Game Theory-Based Multi-Group, Large-Scale Leader–Follower-Distributed Optimal Swarming Control

To achieve the overall swarming behavior for all groups, the overall cost function of the leaders' (Equation (14)) and followers' individual cost function (Equation (15)) need to be minimized. Now, the leader-follower Hamiltonian can be obtained using Bellman's principle of optimality [39] and optimal control [20]. The Hamiltonian of all the leaders is as follows:

$$H_L[x_L, \partial_{x_L} V_L(x_L, \rho)] = \sum_{i=1}^{M} H^i_L[x^i_L, \partial_{x^i_L} V^i_L(x^i_L, x^{-i}_L, \rho^i)] \tag{16}$$

with individual leaders distributed Hamiltonian:

$$H_L^i[x_L^i, \partial_{x_L^i} V_L^i(x_L^i, x_L^{-i}, \rho^i)] = \mathbb{E}\left\{\begin{bmatrix} \Phi_L(x_L^i, x_L^{-i}, u_L^i) + \Phi_{CP,L}(x_L^i, \rho^i) + \partial_{x_L^i} V_L^{i\,T} \\ (x_L^i, x_L^{-i}, \rho^i)[F_a(x_L^i) + G_a(x_L^i)u_L^i] \end{bmatrix}\right\} \tag{17}$$

The Hamiltonian for the follower q from the ith group is derived as follows:

$$H_{F,q}^i[x_{F,q}^i, \partial_{x_{F,q}^i} V_{F,q}^i(x_{F,q}^i, x_L^i, \rho^i)] = \mathbb{E}\left\{\begin{bmatrix} \Phi_F(x_{F,q}^i, x_L^i, u_{F,q}^i) + \Phi_{CA,F}(x_{F,q}^i, \rho^i) + \Phi_{C,F}(x_{F,q}^i, \rho^i) \\ +\partial_{x_{F,q}^i} V_{F,q}^{i\,T}(x_{F,q}^i, x_L^i, \rho^i) + [F_s(x_{F,q}^i) + G_s(x_{F,q}^i)u_{F,q}^i] \end{bmatrix}\right\} \tag{18}$$

Now, the multi-group leaders Hamiltonian–Jacobi–Bellman (HJB) equation can be derived from the cooperative game is as follows:

$$\mathbb{E}\left\{\Phi_{CP,L}(x_L^i, \rho^i)\right\} = \mathbb{E}\left\{\begin{bmatrix} -\partial_t V_L^i(x_L^i, x_L^{-i}, \rho^i) - \sqrt{2\nu^i}\Delta V_L^i(x_L^i, x_L^{-i}, \rho^i) \\ +H_L^i[x_L^i, \partial_{x_L^i} V_L^i(x_L^i, x_L^{-i}, \rho^i)] \end{bmatrix}\right\} \tag{19}$$

Furthermore, a coupled HJB and Fokker–Planck–Kolmogorov(FPK) equations for the large number of followers, using MFG can be obtained as follows:
Follower (HJB):

$$\mathbb{E}\left\{\Phi_{CA,F}(x_{F,q}^i, \rho^i) + \Phi_{C,F}(x_{F,q}^i, \rho^i)\right\} = \mathbb{E}\left\{\begin{bmatrix} -\partial_t V_{F,q}^i(x_{F,q}^i, x_L^i, \rho^i) - \sqrt{2\nu^i}\Delta V_{F,q}^i(x_{F,q}^i, \\ x_L^i, \rho^i) + H_{F,q}^i[x_{F,q}^i, \partial_{x_{F,q}^i} V_{F,q}^i(x_{F,q}^i, x_L^i, \rho^i)] \end{bmatrix}\right\} \tag{20}$$

Follower (FPK):

$$\mathbb{E}\left\{\partial_t\rho^i(x_{F,q}^i, t) - \sqrt{2\nu^i}\Delta\rho^i(x_{F,q}^i, t) - div(\rho^i D_p H_{F,q}^i[x_{F,q}^i, \partial_{x_{F,q}^i} V_{F,q}^i(x_{F,q}^i, x_L^i, \rho^i)])\right\} = 0 \tag{21}$$

Then, according to optimal control theory, the mixed game-based distributed optimal swarming control for a multi-group leader–follower system can be attained as follows:
Leader:

$$u_L^i(x_L^i) = -\frac{1}{2}\mathbb{E}\left\{R_L^{i^{-1}} G_a^T(x_L^i)\partial_{x_L^i} V_L^i(x_L^i, x_L^{-i}, \rho^i)\right\} \tag{22}$$

Follower:

$$u_{F,q}^i(x_{F,q}^i) = -\frac{1}{2}\mathbb{E}\left\{R_{F,q}^{i^{-1}} G_s^T(x_{F,q}^i)\partial_{x_{F,q}^i} V_{F,q}^i(x_{F,q}^i, x_L^i, \rho^i)\right\} \tag{23}$$

Remark 1. *The coupled HJB-FPK equations of the follower and the HJB equation for the leader must be solved in real time in order to determine the optimal control policy. The backward HJB and forward FPK equations, however, are coupled multidimensional nonlinear PDEs that are difficult to solve. Therefore, in this paper, a hierarchical learning-based, multi-actor-critic–mass NN is developed to learn the optimal control online.*

4. Hierarchical Learning-Based Intelligent Optimal Distributed Swarming Control

4.1. Hierarchical Learning-Based Control for Multi-Group Leader–Follower Systems

In this section, a hierarchical learning-based actor–critic–mass algorithm (see Figure 2) is developed and explained in detail. It includes (1) multi-agent actor–critic neural networks (NNs) to obtain the cooperative game-based distributed optimal swarm control for multi-group leaders and (2) actor–critic–mass-based decentralized control for large-scale followers within the same game to obtain the mean field game-based swarming. Additionally, the Stackelberg game integrates distributed swarming control for leaders and decentralized control for followers into a unified framework and further obtains the overall distributed intelligent swarming control for multi-group leader–follower systems.
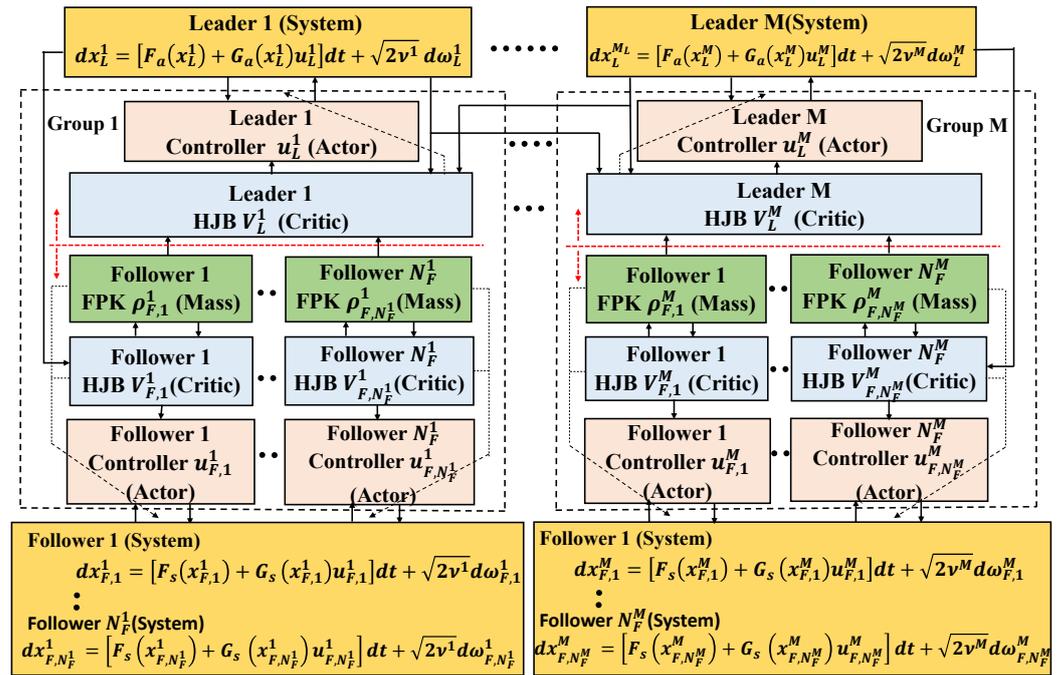
Figure 2. Mixed game-based hierarchical learning structure.

Ideal Actor–Critic Set-up for Leaders:

$$\text{Critic: } V_L^{i*}(x_L^i, x_L^{-i}, \rho^i) = \mathbb{E}\left\{W_{V,L}^{iT}\phi_{V,L}^i + \varepsilon_{\text{HJB}_L}^i\right\}$$
$$\text{Actor: } u_L^{i*}(x_L^i, x_L^{-i}, \rho^i) = \mathbb{E}\left\{W_{u,L}^{iT}\phi_{u,L}^i + \varepsilon_{\text{u,L}}^i\right\}$$

(24)

Ideal Actor–Critic–Mass Set-up for Followers:

$$\text{Critic: } V_{F,q}^{i*}(x_{F,q}^i, x_L^i, \rho^i) = \mathbb{E}\left\{W_{V,F,q}^{iT}\phi_{V,F,q}^i + \varepsilon_{\text{HJB}_{F,q}}^i\right\}$$
$$\text{Actor: } u_{F,q}^{i*}(x_{F,q}^i, x_L^i, \rho^i) = \mathbb{E}\left\{W_{u,F,q}^{iT}\phi_{u,F,q}^i + \varepsilon_{\text{u,F,q}}^i\right\}$$
$$\text{Mass: } \rho^i(x_{F,q}^i, t) = \mathbb{E}\left\{W_{\rho,F,q}^{iT}\phi_{\rho,F,q}^i + \varepsilon_{\text{FPK}_{F,q}}^i\right\}$$

(25)

where $W_{V,L}^i$, $W_{u,L}^i$, $W_{V,F,q}^i$, $W_{u,F,q}^i$, and $W_{\rho,F,q}^i$ are the respective weights for the leader and follower critic, actor, and mass neural network (NN). Also, $\phi(.)$ is the activation function and ε is the reconstruction error of the respective neural networks. Then, the estimated cost and control functions of the leader are as follows:
Estimated Actor-Critic for Multi-Group Leaders:

$$\text{Critic: } \hat{V}_L^i(x_L^i, x_L^{-i}, \bar{\rho}^i) = \mathbb{E}\left\{\hat{W}_{V,L}^{iT}\hat{\phi}_{V,L}^i\right\}$$
$$\text{Actor: } \hat{u}_L^i(x_L^i, x_L^{-i}, \bar{\rho}^i) = \mathbb{E}\left\{\hat{W}_{u,L}^{iT}\hat{\phi}_{u,L}^i\right\}$$

(26)

Please be aware that the leader from each group collects the estimated PDF from each decentralized follower in the same group. Then, the leader evaluates the statistic average of the received PDF (i.e., $\bar{\rho}^i = \sum_{F,q}\hat{\rho}_{F,q}^i$). Next, The residual errors that result from substituting the approximation in Equation (26) into Equation (19) can be utilized to adjust the weights of the leader actor and critic neural networks by using the residual errors given as following equations:

$$\mathbb{E}\left\{e^i_{\text{HJB}_L}\right\} = \mathbb{E}\left\{\left[\Phi_{\text{CP,L}}(x^i_L, \bar{\rho}^i) + \hat{W}^{i^T}_{V,L}[\partial_t \hat{\phi}^i_{V,L} + \sqrt{2\nu^i}\Delta\hat{\phi}^i_{V,L} - \hat{H}^i_{L,W}]\right]\right\} \tag{27}$$

$$\mathbb{E}\left\{e^i_{\text{u,L}}\right\} = \mathbb{E}\left\{\hat{W}^{i^T}_{u,L}\hat{\phi}^i_{u,L} + \frac{1}{2}R^{i^{-1}}_L G^T_a(x^i_L)\partial_{x^i_L}\hat{\phi}^i_{V,L}\right\} \tag{28}$$

where the estimated Hamiltonian of the leader $\hat{H}^i_L$ is defined as

$$\hat{H}^i_L = H^i_L[x^i_L, \partial_{x^i_L}\hat{\phi}^i_{V,L}] = \hat{W}^{i^T}_{V,L}\hat{H}^i_{L,W}$$

Now, let

$$\mathbb{E}\left\{\Psi^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right\} = \mathbb{E}\left\{\partial_t\hat{\phi}_{V,i} + \sqrt{2\nu^i}\Delta\hat{\phi}_{V,i} - \hat{H}^i_{L,W}\right\}$$

and

$$\mathbb{E}\left\{\Phi_{\text{CP,L}}(x^i_L, \bar{\rho}^i)\right\} = \mathbb{E}\left\{\Phi_{\text{CP,L}}(x^i_L, \bar{\rho}^i) + \Phi_{\text{CP,L}}(x^i_L, \rho^i)\right\}$$

Then, the estimation error from the Equation (27) is therefore rewritten follows:

$$\mathbb{E}\left\{e^i_{\text{HJB}_L}\right\} = \mathbb{E}\left\{\left[\Phi_{\text{CP,L}}(x^i_L, \rho^i) + \Phi_{\text{CP,L}}(x^i_L, \bar{\rho}^i) + \hat{W}^{i^T}_{V,L}\Psi^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right]\right\} \tag{29}$$

Next, the effect of the reconstruction error is considered by substituting the optimal function from Equation (24) to the HJB Equation (19):

$$\mathbb{E}\left\{\Phi_{\text{CP,L}}(x^i_L, \rho^i) + W^{i^T}_{V,L}\Psi^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) + \varepsilon^i_{\text{HJB}_L}\right\} = 0 \tag{30}$$

By substituting Equation (30) into Equation (29), the following HJB error equation can be obtained with reconstruction error as:

$$\mathbb{E}\left\{e^i_{\text{HJB}_L}\right\} = \mathbb{E}\left\{\left[\Phi_{\text{CP,L}}(x^i_L, \bar{\rho}^i) - \tilde{W}^{i^T}_{V,L}\hat{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) - W^{i^T}_{V,L}\tilde{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) - \varepsilon^i_{\text{HJB}_L}\right]\right\} \tag{31}$$

Again, the leader actor NN error is as follows:

$$\mathbb{E}\left\{e^i_{\text{u,L}}\right\} = \mathbb{E}\left\{\left[-\tilde{W}^{i^T}_{u,L}\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) - W^{i^T}_{u,L}\tilde{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) - \frac{1}{2}R^{i^{-1}}_L G^T_a(x^i_L)\partial_{x^i_L}\tilde{V}^i_L - \varepsilon^i_{\text{u,L}}\right]\right\} \tag{32}$$

where $\tilde{W}^i_{V,L} = W^i_{V,L} - \hat{W}^i_{V,L}$ and $\tilde{W}^i_{u,L} = W^i_{u,L} - \hat{W}^i_{u,L}$. In addition, the approximated functions can be represented as follows:

$$\mathbb{E}\left\{\tilde{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right\} = \mathbb{E}\left\{\left[\Psi^i_{V,L}(x^i_L, x^{-i}_L, \rho^i) - \hat{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right]\right\}$$

$$\mathbb{E}\left\{\tilde{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right\} = \mathbb{E}\left\{\left[\tilde{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \rho^i) - \hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\right]\right\}$$

Next, the multi-agent actor–critic NNs for the leader are updated by using residual errors (Equations (31) and (32)) with the gradient descent method as follows.
Actor–Critic Update Laws for Multi-Group Leaders:

$$\mathbb{E}\left\{\dot{\hat{W}}^i_{V,L}\right\} = \mathbb{E}\left\{-\alpha^i_{V,L}\frac{\hat{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)e^{i^T}_{\text{HJB}_L}}{1 + \|\hat{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i))\|^2}\right\} \tag{33}$$

$$\mathbb{E}\left\{\dot{\hat{W}}^i_{u,L}\right\} = \mathbb{E}\left\{-\alpha^i_{u,L}\frac{\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)e^{i^T}_{\text{u,L}}}{1 + \|\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\|^2}\right\} \tag{34}$$

Next, the decentralized optimal swarming control for large-scale followers is as follows.

Estimated Actor–Critic Mass for Followers in the Same Group:

$$\text{Critic: } \hat{V}^i_{F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) = \mathbb{E}\left\{ \hat{W}^{i^T}_{V,F,q} \hat{\phi}^i_{V,F,q} \right\}$$

$$\text{Actor: } \hat{u}^i_{F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) = \mathbb{E}\left\{ \hat{W}^{i^T}_{u,F,q} \hat{\phi}^i_{u,F,q} \right\} \tag{35}$$

$$\text{Mass: } \hat{\rho}^i_{F,q}(x^i_{F,q}, t) = \mathbb{E}\left\{ \hat{W}^{i^T}_{\rho,F,q} \hat{\phi}^i_{\rho,F,q} \right\}$$

The residual error after the substitution of (35) into the HJB and FPK equations (Equations (20) and (21)):

$$\mathbb{E}\left\{ e^i_{\text{HJB}_{F,q}} \right\} = \mathbb{E}\left\{ \begin{bmatrix} \Phi_{\text{CA,F}}(x^i_{F,q}, \hat{\rho}^i_{F,q}) + \Phi_{\text{C,F}}(x^i_{F,q}, \hat{\rho}^i_{F,q}) + \hat{W}^{i^T}_{V,F,q} \\ [\partial_t \hat{\phi}^i_{V,F,q} + \sqrt{2\nu^i} \Delta \hat{\phi}^i_{V,F,q} - \hat{H}^i_{F,q,W}] \end{bmatrix} \right\} \tag{36}$$

$$\mathbb{E}\left\{ e^i_{u,F,q} \right\} = \mathbb{E}\left\{ \left[\hat{W}^{i^T}_{u,F,q} \hat{\phi}^i_{u,F,q} + \tfrac{1}{2} R^{i-1}_{F,q} G^T_s(x^i_{F,q}) \partial_{x^i_{F,q}} \hat{\phi}^i_{V,F,q} \right] \right\} \tag{37}$$

$$\mathbb{E}\left\{ e^i_{\text{FPK}_{F,q}} \right\} = \mathbb{E}\left\{ \hat{W}^{i^T}_{\rho,F,q} \left[\partial_t \hat{\phi}^i_{\rho,F,q} - \sqrt{2\nu^i} \Delta \hat{\phi}^i_{\rho,F,q} - div(\hat{\phi}^i_{\rho,F,q}) D_p \hat{H}^i_{F,q} \right] \right\} \tag{38}$$

where

$$\hat{H}^i_{F,q} = H^i_{F,q}[x^i_{F,q}, \partial_{x^i_{F,q}} \hat{\phi}^i_{V,F,q}] = \hat{W}^{i^T}_{V,F,q} \hat{H}^i_{F,q,W}$$

Again, let

$$\mathbb{E}\left\{ \Psi^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) \right\} = \mathbb{E}\left\{ \left[\partial_t \hat{\phi}^i_{V,F,q} + \sqrt{2\nu^i} - \Delta \hat{\phi}^i_{V,F,q} \hat{H}^i_{F,q,W} \right] \right\}$$

$$\mathbb{E}\left\{ \Psi^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t) \right\} = \mathbb{E}\left\{ \left[\partial_t \hat{\phi}^i_{\rho,F,q} - \sqrt{2\nu^i} \Delta \hat{\phi}^i_{\rho,F,q} - div(\hat{\phi}^i_{\rho,F,q}) D_p \hat{H}^i_{F,q} \right] \right\}$$

$$\mathbb{E}\left\{ \Phi_{\text{CA,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) \right\} = \mathbb{E}\left\{ \left[\Phi_{\text{CA,F}}(x^i_{F,q}, \rho^i_{F,q}) - \Phi_{\text{CA,F}}(x^i_{F,q}, \hat{\rho}^i_{F,q}) \right] \right\}$$

$$\mathbb{E}\left\{ \Phi_{\text{C,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) \right\} = \mathbb{E}\left\{ \left[\Phi_{\text{C,F}}(x^i_{F,q}, \rho^i_{F,q}) - \Phi_{\text{C,F}}(x^i_{F,q}, \hat{\rho}^i_{F,q}) \right] \right\}$$

The estimation errors in Equations (36) and (38) can be simplified as follows:

$$\mathbb{E}\left\{ e^i_{\text{HJB}_{F,q}} \right\} = \mathbb{E}\left\{ \begin{bmatrix} \Phi_{\text{CA,F}}(x^i_{F,q}, \rho^i_{F,q}) + \Phi_{\text{CA,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) + \Phi_{\text{C,F}}(x^i_{F,q}, \rho^i_{F,q}) \\ + \Phi_{\text{C,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) + \hat{W}^{i^T}_{V,F,q} \Psi^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) \end{bmatrix} \right\} \tag{39}$$

$$\mathbb{E}\left\{ e^i_{\text{FPK}_{F,q}} \right\} = \mathbb{E}\left\{ \hat{W}^{i^T}_{\rho,F,q} \Psi^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t) \right\} \tag{40}$$

By substituting the optimal functions from Equation (25) into Equations (20) and (21), we obtain

$$\mathbb{E}\left\{ \left[\Phi_{\text{CA,F}}(x^i_{F,q}, \rho^i_{F,q}) + \Phi_{\text{C,F}}(x^i_{F,q}, \rho^i_{F,q}) + W^{i^T}_{V,F,q} \Psi^i_{V,F,q}(x^i_{F,q}, x^i_L, \rho^i_{F,q}) + \varepsilon^i_{\text{HJB}_{F,q}} \right] \right\} = 0 \tag{41}$$

$$\mathbb{E}\left\{ W^{i^T}_{\rho,F,q} \Psi^i_{\rho,F,q}(x^i_{F,q}, V^i_{F,q}, t) + \varepsilon^i_{\text{FPK}_{F,q}} \right\} = 0 \tag{42}$$

By substituting Equations (41) and (42) into Equations (39) and (40), respectively, we obtain

$$\mathbb{E}\left\{ e^i_{\text{HJB}_{F,q}} \right\} = \mathbb{E}\left\{ \begin{bmatrix} \Phi_{\text{CA,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) + \Phi_{\text{C,F}}(x^i_{F,q}, \tilde{\rho}^i_{F,q}) - \tilde{W}^{i^T}_{V,F,q} \hat{\Psi}^i_{V,F,q} \\ (x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) - W^{i^T}_{V,F,q} \Psi^i_{V,F,q}(x^i_{F,q}, x^i_L, \tilde{\rho}^i_{F,q}) - \varepsilon^i_{\text{HJB}_{F,q}} \end{bmatrix} \right\} \tag{43}$$

$$\mathbb{E}\left\{ e^i_{\text{FPK}_{F,q}} \right\} = \mathbb{E}\left\{ \left[-\tilde{W}^{i^T}_{\rho,F,q} \hat{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t) - W^{i^T}_{\rho,F,q} \Psi^i_{\rho,F,q}(x^i_{F,q}, \tilde{V}^i_{F,q}, t) - \varepsilon^i_{\text{FPK}_{F,q}} \right] \right\} \tag{44}$$

Similarly, we can obtain

$$\mathbb{E}\left\{e^i_{u,F,q}\right\} = \mathbb{E}\left\{ \begin{bmatrix} -\tilde{W}^{i^T}_{u,F,q}\hat{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q}) - W^{i^T}_{u,F,q}\tilde{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \tilde{\rho}^i_{F,q}) \\ -\frac{1}{2}R^{i-1}_{F,q}G^T_s(x^i_{F,q})\partial_{x^i_{F,q}}\tilde{V}^i_{F,q} - \varepsilon^i_{u,F,q} \end{bmatrix} \right\} \tag{45}$$

where $\tilde{W}^i_{V,F,q} = W^i_{V,F,q} - \hat{W}^i_{V,F,q}$, $\tilde{W}^i_{u,F,q} = W^i_{u,F,q} - \hat{W}^i_{u,F,q}$ and $\tilde{W}^i_{\rho,F,q} = W^i_{\rho,F,q} - \hat{W}^i_{\rho,F,q}$. In addition, we have

$$\mathbb{E}\left\{\tilde{\Psi}^i_{V,F,q}(x^i_{F,q}, x^i_L, \tilde{\rho}^i_{F,q})\right\} = \mathbb{E}\left\{\left[\Psi^i_{V,F,q}(x^i_{F,q}, x^i_L, \rho^i_{F,q}) - \hat{\Psi}^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})\right]\right\}$$

$$\mathbb{E}\left\{\tilde{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \tilde{\rho}^i_{F,q})\right\} = \mathbb{E}\left\{\left[\phi^i_{u,F,q}(x^i_{F,q}, x^i_L, \rho^i_{F,q}) - \hat{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})\right]\right\}$$

$$\mathbb{E}\left\{\tilde{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \tilde{V}^i_{F,q}, t)\right\} = \mathbb{E}\left\{\left[\Psi^i_{\rho,F,q}(x^i_{F,q}, V^i_{F,q}, t) - \hat{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t)\right]\right\}$$

Using the residual errors in Equations (43)–(45), the actor–critic-mass NNs for the followers are tuned by using gradient descent method as follows.

Actor–Critic–Mass Update Laws for Followers:

$$\mathbb{E}\left\{\hat{W}^i_{V,F,q}\right\} = \mathbb{E}\left\{ -\alpha^i_{V,F,q}\frac{\hat{\Psi}^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})e^{i^T}_{HJB_{F,q}}}{1 + \|\hat{\Psi}^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})\|^2} \right\} \tag{46}$$

$$\mathbb{E}\left\{\hat{W}^i_{u,F,q}\right\} = \mathbb{E}\left\{ -\alpha^i_{u,F,q}\frac{\hat{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})e^{i^T}_{u,F,q}}{1 + \|\hat{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})\|^2} \right\} \tag{47}$$

$$\mathbb{E}\left\{\hat{W}^i_{\rho,F,q}\right\} = \mathbb{E}\left\{ -\alpha^i_{\rho,F,q}\frac{\hat{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t)e^{i^T}_{FPK_{F,q}}}{1 + \|\hat{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t)\|^2} \right\} \tag{48}$$

where $\alpha^i_{V,L}$, $\alpha^i_{u,L}$, $\alpha^i_{V,F,q}$, $\alpha^i_{u,F,q}$, and $\alpha^i_{\rho,F,q}$ are the learning rates:

Remark 2. *The functions* $\hat{\Psi}^i_{V,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)$, $\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)$, $\hat{\Psi}^i_{V,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})$, $\hat{\phi}^i_{u,F,q}(x^i_{F,q}, x^i_L, \hat{\rho}^i_{F,q})$, *and* $\hat{\Psi}^i_{\rho,F,q}(x^i_{F,q}, \hat{V}^i_{F,q}, t)$ *must satisfy the persistent excitation (PE) condition [20] in order to let the weights converge.*

4.2. Optimal Swarming Control Performance Analysis

The performance of all NNs as well as the closed-loop swarming system stability are given in this section:

Theorem 1. *Let* $\mathbb{E}\{\hat{W}^i_{V,L}\}$ *and* $\mathbb{E}\{\hat{W}^i_{V,F,q}\}$ *be updated as Equations (33) and (46) with the learning rates* $\alpha^i_{V,L} > 0$ *and* $\alpha^i_{V,F,q} > 0$, *respectively. Then, the error between the actual and approximated critic NN weights* $\mathbb{E}\{\tilde{W}^i_{V,L}\}$ *and* $\mathbb{E}\{\tilde{W}^i_{V,F,q}\}$ *as well as the optimal evaluation function approximation errors (i.e.,* $\mathbb{E}\{\tilde{V}^i_L\} = \mathbb{E}\{V^i_L - \hat{V}^i_L\}$ *and* $\mathbb{E}\{\tilde{V}^i_{F,q}\} = \mathbb{E}\{V^i_{F,q} - \hat{V}^i_{F,q}\}$*) are uniformly ultimately bounded (UUB). Moreover,* $\mathbb{E}\{\hat{W}^i_{V,L}\}$, $\mathbb{E}\{\tilde{W}^i_{V,F,q}\}$, $\mathbb{E}\{\tilde{V}^i_L\}$, *and* $\mathbb{E}\{\tilde{V}^i_{F,q}\}$ *are asymptotically stable while the reconstruction errors are sufficiently small. Next, the bounds for the evaluation function approximation errors* $\tilde{V}^i_L$ *and* $\tilde{V}^i_{F,q}$ *are as follows:*

$$\mathbb{E}\left\{\|\tilde{V}^i_L(t)\|\right\} = \mathbb{E}\left\{\|\tilde{W}^{i^T}_{V,L}\hat{\phi}^i_{V,L} + W^{i^T}_{V,L}\tilde{\phi}^i_{V,L} + \varepsilon^i_{HJB_L}\|\right\}$$

$$\leq \mathbb{E}\left\{\|\tilde{W}^{i^T}_{V,L}\|\|\hat{\phi}^i_{V,L}\|\right\} + l_{\phi^i_{V,L}}\mathbb{E}\left\{\|W^i_{V,L}\|\|\bar{\rho}^i\|\right\} + \mathbb{E}\left\{\|\varepsilon^i_{HJB_L}\|\right\}$$

$$\leq b^i_{W,V,L}\mathbb{E}\left\{\|\hat{\phi}^i_{V,L}\|\right\} + l_{\phi^i_{V,L}}\mathbb{E}\left\{\|W^i_{V,L}\|\right\}\bar{b}^i_{\rho,F,q} + \mathbb{E}\left\{\|\varepsilon^i_{HJB_L}\|\right\} \equiv b^i_{V,L}(t) \tag{49}$$

Similarly, we have

$$
\begin{aligned}
\mathbb{E}\big\{\|\tilde{V}^i_{F,q}(t)\|\big\} &= \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{V,F,q}\hat{\phi}^i_{V,F,q} + W^{i\,T}_{V,F,q}\tilde{\phi}^i_{V,F,q} + \varepsilon^i_{HJB_{F,q}}\|\big\} \\
&\leq \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{V,F,q}\|\|\hat{\phi}^i_{V,F,q}\|\big\} + l_{\phi^i_{V,F,q}}\mathbb{E}\big\{\|W^i_{V,F,q}\|\|\tilde{\rho}^i_{F,q}\|\big\} + \mathbb{E}\big\{\|\varepsilon^i_{HJB_{F,q}}\|\big\} \\
&\leq b^i_{W,V,F,q}\mathbb{E}\big\{\|\hat{\phi}^i_{V,F,q}\|\big\} + l_{\phi^i_{V,F,q}}\mathbb{E}\big\{\|W^i_{V,F,q}\|\big\}b^i_{\rho,F,q} + \mathbb{E}\big\{\|\varepsilon^i_{HJB_{F,q}}\| \equiv b^i_{V,F,q}(t)
\end{aligned}
\tag{50}
$$

where $l_{\phi^i_{v,L}}$ and $l_{\phi^i_{V,F,q}}$ are the Lipschitz constants of the critic activation functions $\phi^i_{V,L}$ and $\phi^i_{V,F,q'}$ respectively. Additionally, $\bar{b}^i_{\rho,F,q}$ can be calculated by taking the average of the mass bound $b^i_{\rho,F,q}$ of each follower.

Proof. See Appendix A. □

Theorem 2. *Let $\mathbb{E}\{\hat{W}^i_{\rho,F,q}\}$ be updated as in Equation (48), where the learning rate $\alpha^i_{\rho,F,q} > 0$. Then, the error between the actual and approximated mass NN weights $\mathbb{E}\{\tilde{W}^i_{\rho,F,q}\}$ as well as the mass function approximation errors (i.e., $\mathbb{E}\{\tilde{\rho}^i_{F,q}\} = \mathbb{E}\{\rho^i_{F,q} - \hat{\rho}^i_{F,q}\}$) are uniformly ultimately bounded (UUB). Moreover, $\mathbb{E}\{\tilde{W}^i_{\rho,F,q}\}$ and $\mathbb{E}\{\tilde{\rho}^i_{F,q}\}$ are asymptotically stable while the reconstruction errors are sufficiently small. The bound for the mass approximation errors $\tilde{\rho}^i_{F,q}$ is as follows:*

$$
\begin{aligned}
\mathbb{E}\big\{\|\tilde{\rho}^i_{F,q}(t)\|\big\} &= \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{\rho,F,q}\hat{\phi}^i_{\rho,F,q} + \varepsilon^i_{FPK_{F,q}}\|\big\} \\
&\leq \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{\rho,F,q}\|\|\hat{\phi}^i_{\rho,F,q}\|\big\} + \mathbb{E}\big\{\|\varepsilon^i_{FPK_{F,q}}\|\big\} \\
&\leq b^i_{W,\rho,F,q}\mathbb{E}\big\{\|\hat{\phi}^i_{\rho,F,q}\|\big\} + \mathbb{E}\big\{\|\varepsilon^i_{FPK_{F,q}}\| \equiv b^i_{\rho,F,q}(t)
\end{aligned}
\tag{51}
$$

Proof. See Appendix B. □

Theorem 3. *Let $\mathbb{E}\{\hat{W}^i_{u,L}\}$ and $\mathbb{E}\{\hat{W}^i_{u,F,q}\}$ be updated as in Equations (34) and (47), where the learning rates are $\alpha^i_{u,L} > 0$ and $\alpha^i_{u,F,q} > 0$, respectively. Then, the error between the actual and approximated actor NN weights $\mathbb{E}\{\tilde{W}^i_{u,L}\}$ and $\mathbb{E}\{\tilde{W}^i_{u,F,q}\}$ as well as the optimal control approximation errors (i.e., $\mathbb{E}\{\tilde{u}^i_L\} = \mathbb{E}\{u^i_L - \hat{u}^i_L\}$ and $\mathbb{E}\{\tilde{u}^i_{F,q}\} = \mathbb{E}\{u^i_{F,q} - \hat{u}^i_{F,q}\}$, respectively) are uniformly ultimately bounded (UUB). Moreover, $\mathbb{E}\{\tilde{W}^i_{u,L}\}$, $\mathbb{E}\{\tilde{W}^i_{u,F,q}\}$, $\mathbb{E}\{\tilde{u}^i_L\}$, and $\mathbb{E}\{\tilde{u}^i_{F,q}\}$ are asymptotically stable while the reconstruction errors are sufficiently small. Moreover, the bounds for the actor approximation errors $\tilde{u}^i_L$ and $\tilde{u}^i_{F,q}$ are as follows:*

$$
\begin{aligned}
\mathbb{E}\big\{\|\tilde{u}^i_L(t)\|\big\} &= \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{u,L}(t)\hat{\phi}^i_{u,L} + W^{i\,T}_{u,L}\tilde{\phi}^i_{u,L} + \varepsilon^i_{u,L}\|\big\} \\
&\leq \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{u,L}\|\|\hat{\phi}^i_{u,L}\|\big\} + l_{\phi^i_{u,L}}\mathbb{E}\big\{\|W^i_{u,L}\|\|\bar{\tilde{\rho}}^i\|\big\} + \mathbb{E}\big\{\|\varepsilon^i_{u,L}\|\big\} \\
&\leq b^i_{W,u,L}\mathbb{E}\big\{\|\hat{\phi}^i_{u,L}\|\big\} + l_{\phi^i_{u,L}}\mathbb{E}\big\{\|W^i_{u,L}\|\|\big\}\bar{b}^i_{\rho,F,q} + \mathbb{E}\big\{\|\varepsilon^i_{u,L}\|\big\} \equiv b^i_{u,L}(t)
\end{aligned}
\tag{52}
$$

In addition, we have

$$
\begin{aligned}
\mathbb{E}\big\{\|\tilde{u}^i_{F,q}(t)\|\big\} &= \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{u,F,q}\hat{\phi}^i_{u,F,q} + W^{i\,T}_{u,F,q}\tilde{\phi}^i_{u,F,q} + \varepsilon^i_{u,F,q}\|\big\} \\
&\leq \mathbb{E}\big\{\|\tilde{W}^{i\,T}_{u,F,q}\|\|\hat{\phi}^i_{u,F,q}\|\big\} + l_{\phi^i_{u,F,q}}\mathbb{E}\big\{\|W^i_{u,F,q}\|\|\tilde{\rho}^i_{F,q}\|\big\} + \mathbb{E}\big\{\|\varepsilon^i_{u,F,q}\|\big\} \\
&\leq b^i_{W,u,F,q}\mathbb{E}\big\{\|\hat{\phi}^i_{u,F,q}\|\big\} + l_{\phi^i_{u,F,q}}\mathbb{E}\big\{\|W^i_{u,F,q}\|\big\}b^i_{\rho,F,q}\mathbb{E}\big\{\|\varepsilon^i_{u,F,q}\|\big\} \equiv b^i_{u,F,q}(t)
\end{aligned}
\tag{53}
$$

where $l_{\phi^i_{u,L}}$ and $l_{\phi^i_{u,F,q}}$ are the Lipschitz constants of the actor activation functions $\phi^i_{u,L}$ and $\phi^i_{u,F,q'}$ respectively.

Proof. See Appendix C. □

Next, the closed-loop stability of a multi-group, large-scale leader–follower swarming system with developed hierarchical learning control is analyzed:

Lemma 1. *With the given stochastic error dynamics in Equations (4) and (10), there exist optimal control policies for leaders and followers u_L^i and $u_{F,q}^i$ which satisfy*

$$\mathbb{E}\left\{ e_L^{iT}\left[F_{ar}(e_L^i) + G_{ar}(e_L^i)u_L^i + \frac{\sqrt{2\nu^i}d\omega_L^i}{dt}\right]\right\} \leq -\gamma_1 \mathbb{E}\left\{ \|e_L^i\|^2\right\} \tag{54}$$

$$\mathbb{E}\left\{ e_{F,q}^{iT}\left[F_{sr}(e_{F,q}^i) + G_{sr}(e_{F,q}^i)u_{F,q}^i + \frac{\sqrt{2\nu^i}d\omega_{F,q}^i}{dt}\right]\right\} \leq -\gamma_2 \mathbb{E}\left\{ \|e_{F,q}^i\|^2\right\} \tag{55}$$

Theorem 4 (Closed-Loop Stability). *Let the leaders' and followers' critic, mass, and actor NN weights be updated as in Equations (33)–(48). Additionally, consider the learning rates $\alpha_{V,L}^i$, $\alpha_{V,F,q}^i$, $\alpha_{u,L}^i$, $\alpha_{u,F,q}^i$, and $\alpha_{\rho,F,q}^i$ to be greater than zero. Then, $\mathbb{E}\{\tilde{W}_{V,L}^i\}$, $\mathbb{E}\{\tilde{W}_{V,F,q}^i\}$, $\mathbb{E}\{\tilde{W}_{u,L}^i\}$, $\mathbb{E}\{\tilde{W}_{u,F,q}^i\}$, $\mathbb{E}\{\tilde{W}_{\rho,F,q}^i\}$, $\mathbb{E}\{\tilde{V}_L^i\}$, $\mathbb{E}\{\tilde{V}_{F,q}^i\}$, $\mathbb{E}\{\tilde{u}_L^i\}$, $\mathbb{E}\{\tilde{u}_{F,q}^i\}$, $\mathbb{E}\{\tilde{\rho}_{F,q}^i\}$, $\mathbb{E}\{e_L^i\}$, and $\mathbb{E}\{e_{F,q}^i\}$ are all UUB. Moreover, $\mathbb{E}\{\tilde{W}_{V,L}^i\}$, $\mathbb{E}\{\tilde{W}_{V,F,q}^i\}$, $\mathbb{E}\{\tilde{W}_{u,L}^i\}$, $\mathbb{E}\{\tilde{W}_{u,F,q}^i\}$, $\mathbb{E}\{\tilde{W}_{\rho,F,q}^i\}$, $\mathbb{E}\{\tilde{V}_L^i\}$, $\mathbb{E}\{\tilde{V}_{F,q}^i\}$, $\mathbb{E}\{\tilde{u}_L^i\}$, $\mathbb{E}\{\tilde{u}_{F,q}^i\}$, $\mathbb{E}\{\tilde{\rho}_{F,q}^i\}$, $\mathbb{E}\{e_L^i\}$, and $\mathbb{E}\{e_{F,q}^i\}$ are asymptotically stable while the reconstruction errors are sufficiently small.*

Proof. See Appendix D. □

5. Simulation Results

The developed mixed game theory and hierarchical learning-based intelligent distributed swarming control algorithm is implemented in a very large-scale unmanned aerial vehicle (UAV) system. The experiment aims to validate the effectiveness of the swarming behavior of multiple UAV groups by using the developed techniques.

5.1. Performance Evaluation of Mixed Game Theory-Based Intelligent Distributed Swarm Control

Let, there are four groups in an area scaled to 20×10.

Each group had 1 leader and 500 followers. Each leader was given a predefined time-varying trajectory. Then, to produce the swarming behavior, each leader is tracked by his respective followers. Please note that the leader's location is known to his corresponding followers.

Next, the selected initial positions of the leaders were $x_L^1 = \begin{bmatrix} 2.5 & 6.6 \end{bmatrix}^T$, $x_L^2 = \begin{bmatrix} 2 & 6.2 \end{bmatrix}^T$, $x_L^3 = \begin{bmatrix} 2.5 & 5.8 \end{bmatrix}^T$, and $x_L^4 = \begin{bmatrix} 2.2 & 5.4 \end{bmatrix}^T$. Moreover, the initial state of the followers from each group was generated using the following normal distribution. Group 1: $\mathcal{N}(\mu = [2.5, 6.6]^T, \sigma = 0.4 \times I_2)$; Group 2: $\mathcal{N}(\mu = [2, 6.2]^T, \sigma = 0.4 \times I_2)$; Group 3: $\mathcal{N}(\mu = [2.5, 5.8]^T, \sigma = 0.4 \times I_2)$; and Group 4: $\mathcal{N}(\mu = [2.2, 5.4]^T, \sigma = 0.4 \times I_2)$.

Next, the time-varying reference trajectory was defined as follows:

$$x_d(t) = \begin{bmatrix} 2.9t + 5 & 0.01\sin(2.5t) + 6 \end{bmatrix}^T$$

Additionally, the intrinsic dynamics of the leaders were selected as follows:

$$F_a(x_L^i) = \begin{bmatrix} -x_{L,1}^i + \frac{1}{2}x_{L,2}^{i2} \\ -0.4x_{L,2}^{i2} \end{bmatrix}, \quad G_a(x_L^i) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

with $x_L^i = \begin{bmatrix} x_{L,1}^i & x_{L,2}^i \end{bmatrix}^T$.

Similarly, the follower's dynamics could be derived as follows:

$$F_s(x^i_{F,q}) = \begin{bmatrix} -x^i_{F,q,1} + \frac{1}{2}x^{i2}_{F,q,2} \\ -0.2x^{i2}_{F,q,2} \end{bmatrix}, \quad G_s(x^i_{F,q}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

with $x^i_{F,q} = \begin{bmatrix} x^i_{F,q,1} & x^i_{F,q,2} \end{bmatrix}^T$.

Furthermore, the parameters for evaluating the cost functions (Equations (14) and (15)) were defined as $\nu^i = 0.02$, $h = 1.5$, $r^i = r^j = r = 0.5$, $R^i_L = R^i_{F,q} = 5$, $w_{CA,L} = 5$, $w_{CP,L} = 5$, $w_{CP,F} = 5$, $w_{CA,F} = 5$, and $w_{C,F} = 5$. The total simulation time for this experiment was 15 s.

Next, hierarchical learning-based multi-agent actor–critic neural networks for the leaders and actor–critic–mass neural networks for the followers were constructed. The neural networks learning rate parameters were selected as $\alpha^i_{V,L} = 2 \times 10^{-6}$, $\alpha^i_{V,F,q} = 2 \times 10^{-5}$, $\alpha^i_{u,L} = 2 \times 10^{-4}$, $\alpha^i_{u,F,q} = 2 \times 10^{-3}$, and $\alpha^i_{\rho,F,q} = 1 \times 10^{-3}$.

The trajectories of the multi-group, large-scale UAVs are plotted in Figure 3 for times $t = 0$ s, $t = 5$ s, $t = 10$ s, and $t = 15$ s. The red curve represents the reference trajectory for all groups. Moreover, the leader trajectories are shown with green curves in the figure. For the follower trajectories, different colors have been used. Figure 3a shows the initial positions of all the leaders and followers and the reference trajectory. Then, the leaders and followers from all groups begin their motion from the left region and move toward the right region. The trajectories of the UAVs with swarming behavior are shown in Figure 3b–d. These figures clearly show that all the leaders tracked the reference trajectory and the followers tracked their respective leaders while avoiding collisions to achieve the collective swarming behavior.
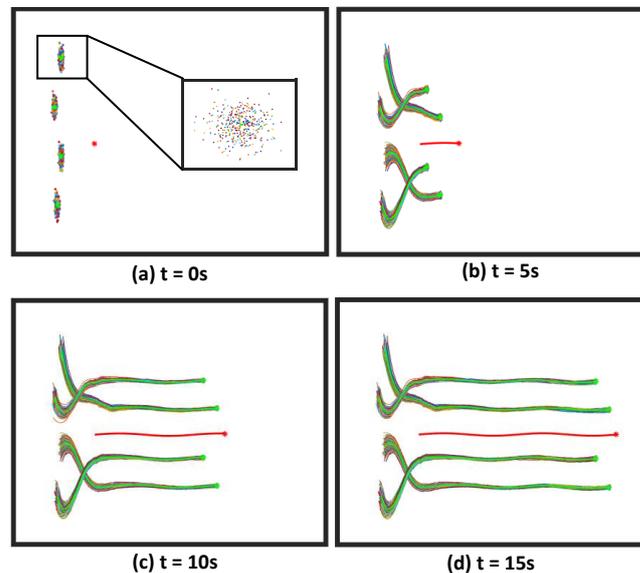

(a) t = 0s

(b) t = 5s

(c) t = 10s

(d) t = 15s

Figure 3. Large-scale leader–follower multi-group swarming. The reference trajectory is denoted by a red curve, the green curve represents the leader trajectory, and the Followers are represented with multiple colors.

The tracking performance of the leaders and followers from all the groups is verified in Figures 4 and 5. Figure 4 shows the tracking error of the leaders from all four groups. From this figure, it is clear that the error converged to zero after a certain time period. This implies that the leaders attained the desired swarming behavior as time progressed. We also needed to ensure that each follower could track their respective leaders to achieve the group swarming behavior. In this regard, coupling functions $\Phi_{CP,L}$ and $\Phi_{CP,F}$ were used in the leader and follower cost functions to ensure that the followers could stay close to their respective leaders. To verify the performance of the developed method, the tracking

errors of the followers with respect to their corresponding leaders are plotted in Figure 5. The average distance of all the followers with respect to their leaders from each group was calculated and plotted in Figure 5a. This figure implies that the tracking errors of the followers from all groups converged to zero over time. To demonstrate the performance clearly, the PDF of the follower tracking error is shown in Figure 5b. The yellow color from time 5 s to 15 s shows the tracking error with higher probability.
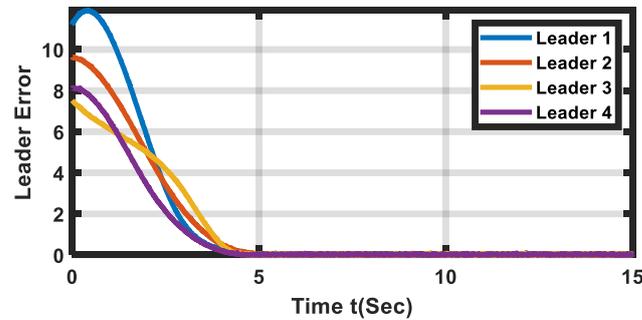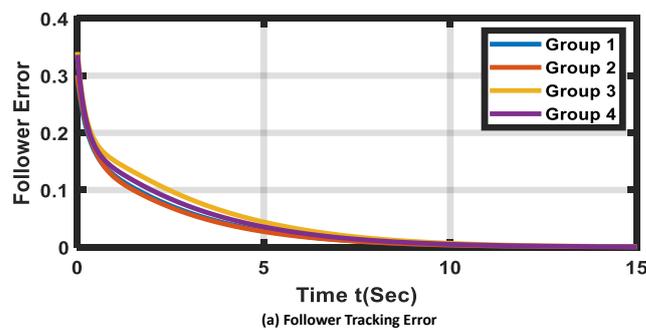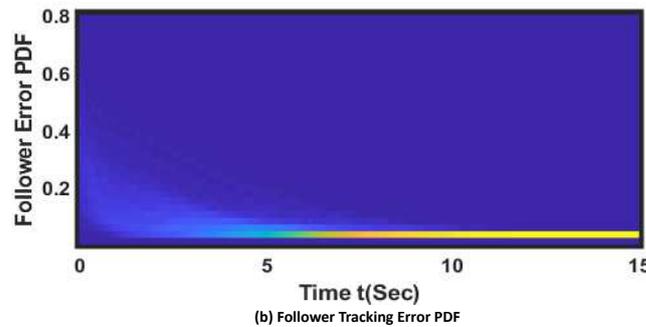

Figure 4. The tracking errors for the leaders.

(a) Follower Tracking Error

(b) Follower Tracking Error PDF

Figure 5. The tracking errors and the tracking error PDF of the followers.

The neural network performance was evaluated by demonstrating the HJB equation error of the leader and the HJB and FPK equation errors of the followers. The HJB errors of the leader and the follower 1 in group 1 have been shown in this simulation. In Figure 6, the HJB equation error of the leader of group 1 is plotted. In the figure, it is clear that the error converged to zero with time. A small window is also demonstrated in this figure for time 14.5–15 s to observe the performance in detail. This implies the optimality of the leader for achieving the desired swarming behavior. Next, the HJB equation error of follower 1 from group 1 is demonstrated in Figure 7. We can clearly see that the HJB error of the followers converged to 0 after 2.5 s. These two figures confirm the optimality of all groups. To confirm the convergence of the mean-field swarming error, the FPK equation error for group 1 follower 1 is presented in Figure 8. The FPK error figure validates the follower mean field equation solution.
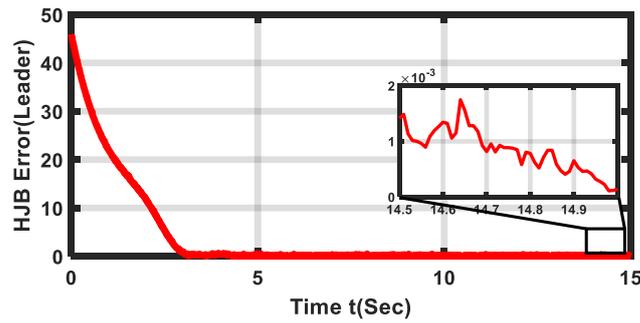
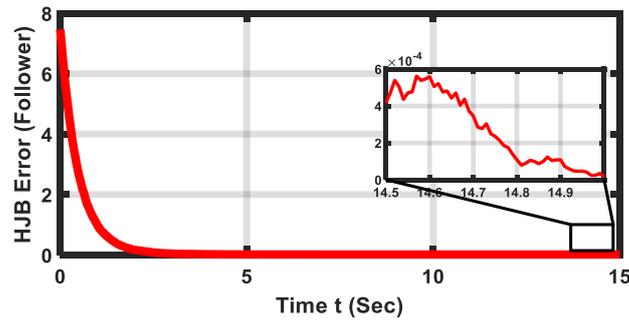**Figure 6.** HJB equation error of the leader in group 1.

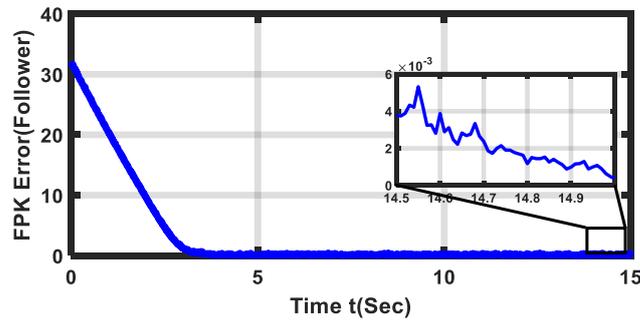**Figure 7.** HJB equation error of follower 1 in group 1.

Figure 8. FPK equation error of follower 1 in group 1.

5.2. Performance Comparison of Mixed Game Theory against Traditional Cooperative Control

Finally, the performance of the developed mixed game theory-based distributed control was compared with the traditional cooperative centralized control [33–35] to demonstrate the significance of our method. In this comparison, each group in the cooperative swarm control scheme had 1 leader and 100 followers. The other parameter values used in the cooperative swarm control scheme (e.g., the initial positions) were identical to those used for the mixed game theory-based distributed swarm control scheme. Next, the running cost including the communication costs was evaluated to provide a comparison of the performances. The running cost of the leader for this simulation was defined as

$$J_L^i = \mathbb{E}\left\{ \int_0^\infty \left[\Phi_L(x_L^i, x_L^{-i}, u_L^i) + \Phi_{CP,L}(x_L^i, \bar{x}^i) + w_c N_F^i + w_{c,L}(M-1) \right] dt \right\} \tag{56}$$

where $\Phi_L(x_L^i, x_L^{-i}, u_L^i) = \Phi_{SE,L}(x_L^i) + \Phi_{CA,L}(x_L^i, x_L^{-i}) + \|u_L^i\|_{R_L^i}^2$ and $w_c N_F^i$ is a new term which represents the communication cost of the leader with their followers from group i. Additionally, w_c is the communication cost weight, and N_F^i is the total number of followers in group i, while $w_{c,L}(M-1)$ represents the communication cost of the leader with other leaders from different groups. Here, $w_{c,L}$ is the communication cost weight, and M is the total number of leaders.

Figure 9 shows the performance of the developed algorithm in terms of the running cost for the leader in group 1. From Figure 9, it is clear that the performance of our approach against the cooperative multi-agent centralized approach was better after a certain amount of time. In the beginning, the cost of the cooperative game-based optimal solution was lower. However, the developed algorithm slowly outperformed the cooperative game-based algorithm. The main reason for this is that the cost function of cooperative centralized swarm control is penalized for high amounts of communication between the leader and the followers in the same group. Similarly, the performance of the developed mixed game theory-based distributed swarm control algorithm for the follower in the same group is shown.
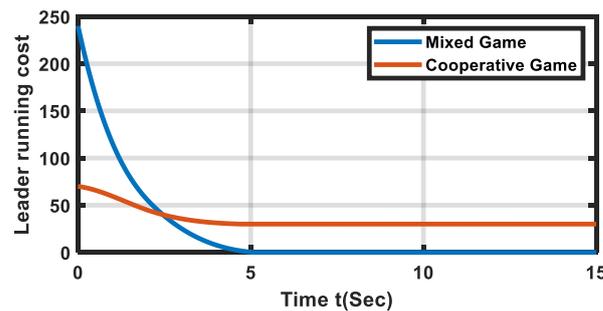

Figure 9. Leader running cost for developed mixed game theory-based distributed approach and the traditional cooperative game-based centralized approach.

The running cost of the follower is defined as

$$J_{F,q}^i = \mathbb{E}\left\{ \int_0^\infty \left[\Phi_F(x_{F,q}^i, x_L^i, u_{F,q}^i) + \Phi_{CA,F}(x_{F,q}^i, \bar{x}^i) + \Phi_{C,F}(x_{F,q}^i, \bar{x}^i) + w_{c,F} N_F^i \right] dt \right\} \quad (57)$$

$$\text{s.t. } \| x_{F,q}^{i,o} - x_L^i \| < r^i \quad (58)$$

with $\Phi_F(x_{F,q}^i, x_L^i, u_{F,q}^i) = \Phi_{SE,F}(x_{F,q}^i) + \Phi_{CP,F}(x_{F,q}^i, x_L^i) + \| u_{F,q}^i \|_{R_{F,q}^i}^2$, $w_{c,F} N_F^i$ being a new term which represents the communication cost of follower q from group i. In addition, $w_{c,F}$ is the communication cost weight, and N_F^i is the total number of followers in group i. Similar to Figure 9, the performance of the developed distributed swarm algorithm in terms of the running costs for the followers in group 1 are demonstrated in Figure 10. The cost of the cooperative centralized approach was penalized here because of the communication among a large number of followers in the same group. From Figures 9 and 10, it is clear that the developed algorithm outperformed the traditional cooperative centralized approach.
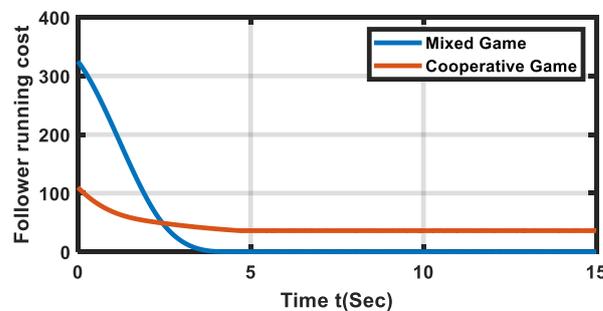

Figure 10. Follower running cost for developed mixed game theory-based distributed approach and the traditional cooperative game-based centralized approach.

6. Conclusions

This paper developed mixed game-based distributed intelligent swarming control along with a hierarchical learning algorithm for multi-group, large-scale leader–follower

systems. To attain the collective swarming behavior for a large number of agents, a mixed game theory was designed which included a cooperative game to ensure collective swarming behavior among the multi-group leaders, a Stackelberg game to bond the group leader with its large-scale followers, and a mean field game to continue the collective swarming behavior to all the followers without raising computational complexity by breaking the "Curse of Dimensionality". Moreover, a hierarchical learning-based actor–critic algorithm was designed to achieve the solution of intelligent optimal swarming control. This structure includes the multi-agent actor–critic neural networks to learn the distributed swarm control for multi-group leaders and actor–critic–mass-based neural networks to learn decentralized swarm control for the large-scale followers. The developed mixed game-based intelligent swarm control optimizes the collective swarming behavior and also adapts to uncertainties in the dynamic environment. Finally, the effectiveness of the developed techniques was validated through Lyapunov analysis and numerical simulations.

Author Contributions: Conceptualization, S.D. and H.X.; methodology, S.D. and H.X.; software, S.D.; validation, S.D. and H.X.; formal analysis, S.D. and H.X.; investigation, H.X.; resources, H.X.; data curation, S.D.; writing—original draft preparation, S.D.; writing—review and editing, H.X.; visualization, S.D.; supervision, H.X.; project administration, H.X.; funding acquisition, H.X. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 1

Leader Critic NN: Consider the following Lyapunov function candidate:

$$L_{V,L}^i(t) = \frac{1}{2}\mathrm{tr}\left\{\mathbb{E}\{\tilde{W}_{V,L}^{iT}\tilde{W}_{V,L}^i\}\right\} \tag{A1}$$

In addition, the first derivative of the leader–critic NN estimated weight from Equation (33) can be obtained as follows:

$$\mathbb{E}\{\dot{\tilde{W}}_{V,L}^i\} = \mathbb{E}\{-\dot{\hat{W}}_{V,L}^i\} = \mathbb{E}\left\{\alpha_{V,L}^i\frac{\hat{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i)e_{\mathrm{HJB_L}}^{iT}}{1 + \|\hat{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i))\|^2}\right\} \tag{A2}$$

According to the Lyapunov stability analysis, we take the first derivative of Equation (A1) and substitute the leader–critic NN weight estimation error dynamic from Equation (A2):

$$\dot{L}_{V,L}^i(t) = \mathrm{tr}\left\{\mathbb{E}\{\tilde{W}_{V,L}^{iT}\dot{\tilde{W}}_{V,L}^i\}\right\} = \alpha_{V,L}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^{iT}\frac{\hat{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i)e_{\mathrm{HJB_L}}^{iT}}{1 + \|\hat{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i))\|^2}\right\}\right) \tag{A3}$$

Then, we let

$$\hat{\Psi}_{V,L}^i = \hat{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i) \; ; \; \breve{\Psi}_{V,L}^i = \breve{\Psi}_{V,L}^i(x_L^i, x_L^{-i}, \bar{\rho}^i)$$

Now, by substituting Equation (31) into Equation (A3), we can obtain

$$\dot{L}_{V,L}^i(t) \le \alpha_{V,L}^i\mathrm{tr}\left(\mathbb{E}\left\{\left\{\frac{\tilde{W}_{V,L}^{iT}\hat{\Psi}_{V,L}^i}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\}\left\{\tilde{\Phi}_{\mathrm{CP,L}} - \tilde{W}_{V,L}^{iT}\hat{\Psi}_{V,L}^i - W_{V,L}^{iT}\breve{\Psi}_{V,L}^i - \varepsilon_{\mathrm{HJB_L}}^i\right\}^T\right\}\right)$$

$$\le \alpha_{V,L}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^{iT}\frac{\hat{\Psi}_{V,L}^i\tilde{\Phi}_{\mathrm{CP,L}}^T}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\}\right) - \alpha_{V,L}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^i\frac{\hat{\Psi}_{V,L}^i\hat{\Psi}_{V,L}^{iT}\tilde{W}_{V,L}^i}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\}\right) - \alpha_{V,L}^i\mathrm{tr}$$

$$\left(\mathbb{E}\left\{\tilde{W}_{V,L}^{i^T}\frac{\hat{\Psi}_{V,L}^i\tilde{\Psi}_{V,L}^{i^T}W_{V,L}^i}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right)-\alpha_{V,L}^i\operatorname{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^i\frac{\hat{\Psi}_{V,L}^i\varepsilon_{\text{HJB}_\text{L}}^{i^T}}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right) \tag{A4}$$

Next, the triangle inequality properties are applied to Equation (A4):

$$\dot{L}_{V,L}^i(t)\leq-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\operatorname{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^i\right.\right.$$

$$\left.\left.\times\frac{\hat{\Psi}_{V,L}^i\tilde{\Phi}_{\text{CP,L}}^T}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right)-\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{\text{CP,L}}\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{\text{CP,L}}\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\frac{1}{4}\alpha_{V,L}^i$$

$$\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\operatorname{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,L}^{i^T}\frac{\hat{\Psi}_{V,L}^i\tilde{\Psi}_{V,L}^{i^T}W_{V,L}^i}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right)-\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}$$

$$+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\operatorname{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,i}^T\frac{\hat{\Psi}_{V,L}^i\varepsilon_{\text{HJB}_\text{L}}^{i^T}}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right)-$$

$$\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\} \tag{A5}$$

Now, Equation (A5) can be simplified using $\frac{1}{4}a^2\pm ab+b^2=\left(\frac{1}{2}a\pm b\right)^2$:

$$\dot{L}_{V,L}^i(t)\leq-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\frac{1}{1+\|\hat{\Psi}_{V,L}^i\|^2}\left[\frac{1}{4}\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2-\operatorname{tr}\right.\right.$$

$$\left.\left.\{\hat{\Psi}_{V,L}^i\tilde{W}_{V,L}^{i^T}\tilde{\Phi}_{\text{CP,L}}^T\}+\|\tilde{\Phi}_{\text{CP,L}}\|^2\right]\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\frac{1}{1+\|\hat{\Psi}_{V,L}^i\|^2}\left[\frac{1}{4}\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2+\operatorname{tr}\{\tilde{W}_{V,L}^{i^T}\hat{\Psi}_{V,L}^i\right.\right.$$

$$\left.\left.\tilde{\Psi}_{V,L}^{i^T}W_{V,L}^i\}+\|W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2\right]\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\frac{1}{1+\|\hat{\Psi}_{V,L}^i\|^2}\left[\frac{1}{4}\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2+\tilde{W}_{V,L}^{i^T}\hat{\Psi}_{V,L}^i\varepsilon_{\text{HJB}_\text{L}}^{i^T}\right.\right.$$

$$\left.\left.+\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2\right]\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{\text{CP,L}}\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}$$

$$\leq-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\frac{\tilde{W}_{V,L}^{i^T}\hat{\Psi}_{V,L}^i}{2}-\tilde{\Phi}_{\text{CP,L}}\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\right.$$

$$\left.\frac{\|\frac{\tilde{W}_{V,L}^{i^T}\hat{\Psi}_{V,L}^i}{2}+W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}-\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\frac{\tilde{W}_{V,L}^{i^T}\hat{\Psi}_{V,L}^i}{2}+\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{\text{CP,L}}\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}$$

$$+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|W_{V,L}^{i^T}\tilde{\Psi}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,i}^i\|^2}\right\} \tag{A6}$$

By dropping several negative terms, Equation (A6) can be simplified as follows:

$$\dot{L}_{V,L}^i(t)\leq-\frac{1}{4}\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2\|\tilde{W}_{V,L}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}+B_{W,V,L}^i(t) \tag{A7}$$

where

$$B_{W,V,L}^i(t)=\left[\alpha_{V,L}^i\frac{[l_{\Phi_{\text{CP,L}}^i}+l_{\Psi_{V,L}^i}\mathbb{E}\{\|W_{V,L}^i\|^2\}\|\bar{\rho}^i\|^2\}}{1+\mathbb{E}\{\|\hat{\Psi}_{V,L}^i\|^2\}}+\alpha_{V,L}^i\mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_\text{L}}^i\|^2}{1+\|\hat{\Psi}_{V,L}^i\|^2}\right\}\right] \tag{A8}$$

Here, $l_{\Phi_{\text{CP,L}}^i}$ and $l_{\Psi_{V,L}^i}$ are the Lipschitz constants, and $\bar{\rho}^i$ is a mass estimation bound.

The leader–critic NN weight estimation error will be UUB, and the bound is

$$\mathbb{E}\left\{\|\tilde{W}_{V,L}^i\|\right\} \le 2\mathbb{E}\left\{\sqrt{\frac{(1+\|\hat{\Psi}_{V,L}^i\|^2)}{\alpha_{V,L}^i\|\hat{\Psi}_{V,L}^i\|^2}B_{W,V,L}^i}\right\} \equiv b_{W,V,L}^i \tag{A9}$$

This completes the proof.

Follower Critic NN: Consider the following Lyapunov function candidate:

$$L_{V,F,q}^i(t) = \frac{1}{2}\mathrm{tr}\left\{\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\tilde{W}_{V,F,q}^i\right\}\right\} \tag{A10}$$

Additionally, the first derivative of the leader–critic NN estimated weight from Equation (47) can be obtained as follows:

$$\mathbb{E}\left\{\dot{\tilde{W}}_{V,F,q}^i\right\} = \mathbb{E}\left\{-\dot{\hat{W}}_{V,F,q}^i\right\} = \mathbb{E}\left\{\alpha_{V,F,q}^i\frac{\hat{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)e_{\mathrm{HJB}_{F,q}}^{iT}}{1+\|\hat{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)\|^2}\right\} \tag{A11}$$

According to Lyapunov stability analysis, we take the first derivative of Equation (A19) and substitute the follower–critic NN weight estimation error dynamic from Equation (A11):

$$\dot{L}_{V,F,q}^i(t) = \mathrm{tr}\left\{\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\dot{\tilde{W}}_{V,F,q}^i\right\}\right\} = \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)e_{\mathrm{HJB}_{F,q}}^{iT}}{1+\|\hat{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)\|^2}\right\}\right) \tag{A12}$$

Then, we let

$$\hat{\Psi}_{V,F,q}^i = \hat{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)\ ;\ \check{\Psi}_{V,F,q}^i = \check{\Psi}_{V,F,q}^i(x_{F,q}^i, x_L^i, \hat{\rho}_{F,q}^i)$$

Now, by substituting Equation (43) into Equation (A12), we can obtain

$$\dot{L}_{V,F,q}^i(t) \le \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\left\{\frac{\tilde{W}_{V,F,q}^{iT}\hat{\Psi}_{V,F,q}^i}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\left\{\check{\Phi}_{\mathrm{CA,F}} + \check{\Phi}_{\mathrm{C,F}} - \tilde{W}_{V,F,q}^{iT}\hat{\Psi}_{V,F,q}^i - W_{V,F,q}^{iT}\check{\Psi}_{V,F,q}^i\right.\right.\right.$$

$$\left.\left.\left. - \varepsilon_{\mathrm{HJB}_{F,q}}^i\right\}^T\right\}\right) \le \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i\check{\Phi}_{\mathrm{CA,F}}^T}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right) + \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i\check{\Phi}_{\mathrm{C,F}}^T}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right)$$

$$ - \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^i\frac{\hat{\Psi}_{V,F,q}^i\hat{\Psi}_{V,F,q}^{iT}\tilde{W}_{V,F,q}^i}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right) - \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i\check{\Psi}_{V,F,q}^{iT}W_{V,F,q}^i}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right)$$

$$ - \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^i\frac{\hat{\Psi}_{V,F,q}^i\varepsilon_{\mathrm{HJB}_{F,q}}^{iT}}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right) \tag{A13}$$

Next, the triangle inequality properties are applied to Equation (A13):

$$\dot{L}_{V,F,q}^i(t) \le -\frac{1}{4}\alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2\|\tilde{W}_{V,F,q}^i\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\} + \alpha_{V,F,q}^i\mathrm{tr}\times\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i\check{\Phi}_{\mathrm{CA,F}}^T}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right)$$

$$ - \alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\check{\Phi}_{\mathrm{CA,F}}\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\} + \alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\check{\Phi}_{\mathrm{CA,F}}\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\} - \frac{1}{4}\alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2\|\tilde{W}_{V,F,q}^i\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}$$

$$ + \alpha_{V,F,q}^i\mathrm{tr}\left(\mathbb{E}\left\{\tilde{W}_{V,F,q}^{iT}\frac{\hat{\Psi}_{V,F,q}^i\check{\Phi}_{\mathrm{C,F}}^T}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}\right) - \alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\check{\Phi}_{\mathrm{C,F}}\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\} + \alpha_{V,F,q}^i\mathbb{E}\left\{\frac{\|\check{\Phi}_{\mathrm{C,F}}\|^2}{1+\|\hat{\Psi}_{V,F,q}^i\|^2}\right\}$$

$$- \frac{1}{4} \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}_{V,F,q}^{i^T} \frac{\hat{\Psi}_{V,F,q}^{i} \hat{\Psi}_{V,F,q}^{i^T} W_{V,F,q}^{i}}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} \right) - \alpha_{V,F,q}^{i}$$

$$\mathbb{E}\left\{ \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \frac{1}{4}\alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} -$$

$$\alpha_{V,F,q}^{i} \mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}_{V,F,q}^{i^T} \frac{\hat{\Psi}_{V,F,q}^{i} \varepsilon_{\mathrm{HJB}_{F,q}}^{i^T}}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} \right) - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} \tag{A14}$$

Now, Equation (A14) can be simplified as follows:

$$\dot{L}_{V,F,q}^{i}(t) \leq -\alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{1}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \left[\frac{1}{4} \|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2 - \mathrm{tr}\{\hat{\Psi}_{V,F,q}^{i} \tilde{W}_{V,F,q}^{i^T} \tilde{\Phi}_{\mathrm{CA,F}}^{T}\} + \| \right.\right.$$

$$\left.\left. \tilde{\Phi}_{\mathrm{CA,F}}\|^2 \right] \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{1}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \left[\frac{1}{4} \|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2 - \mathrm{tr}\{\hat{\Psi}_{V,F,q}^{i} \tilde{W}_{V,F,q}^{i^T} \tilde{\Phi}_{\mathrm{C,F}}^{T}\} + \|\tilde{\Phi}_{\mathrm{C,F}}\|^2 \right.\right.$$

$$\left.\left. \right] \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{1}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \left[\frac{1}{4} \|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2 + \mathrm{tr}\{\tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i} \tilde{\Psi}_{V,F,q}^{i^T} W_{V,F,q}^{i}\} + \|W_{V,F,q}^{i^T} \right.\right.$$

$$\left.\left. \tilde{\Psi}_{V,F,q}^{i}\|^2 \right] \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{1}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \left[\frac{1}{4} \|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2 + \tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i} \varepsilon_{\mathrm{HJB}_{F,q}}^{i} + \right.\right.$$

$$\left.\left. \|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2 \right] \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{CA,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{C,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}$$

$$\left\{ \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\}$$

$$\leq -\alpha_{V,F,q}^{i} \mathbb{E} \times \left\{ \frac{\|\frac{\tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} - \tilde{\Phi}_{\mathrm{CA,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\frac{\tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} - \tilde{\Phi}_{\mathrm{C,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \right.$$

$$\frac{\|\frac{\tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} + W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\frac{\tilde{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} + \varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{CA,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\}$$

$$+ \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{C,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\}$$

$$\leq -\alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\frac{\bar{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} - \tilde{\Phi}_{\mathrm{CA,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\frac{\bar{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} - \tilde{\Phi}_{\mathrm{C,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \mathbb{E}\left\{ \frac{\alpha_{V,F,q}^{i}}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right.$$

$$\times \left[\begin{matrix} \frac{\|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\bar{W}_{V,F,q}^{i}\|^2}{4} + \\ \frac{\|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\bar{W}_{V,F,q}^{i}\|^2}{2} \\ + \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{2} \\ + \|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2 \end{matrix} \right] \right\} - \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\frac{\bar{W}_{V,F,q}^{i^T} \hat{\Psi}_{V,F,q}^{i}}{2} + \varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{CA,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\}$$

$$+ \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\tilde{\Phi}_{\mathrm{C,F}}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} + \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|W_{V,F,q}^{i^T} \tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} \alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\varepsilon_{\mathrm{HJB}_{F,q}}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\}$$

$$\leq -\frac{3}{4}\alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|\hat{\Psi}_{V,F,q}^{i}\|^2 \|\tilde{W}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \frac{3}{2}\alpha_{V,F,q}^{i} \mathbb{E}\left\{ \frac{\|W_{V,F,q}^{i^T}\|^2 \|\tilde{\Psi}_{V,F,q}^{i}\|^2}{1 + \|\hat{\Psi}_{V,F,q}^{i}\|^2} \right\} - \alpha_{V,F,q}^{i} \mathbb{E}$$

$$\left\{\frac{\|\frac{\tilde{W}_{V,F,q}^{i^{T}}\hat{\Psi}_{V,F,q}^{i}}{2}-\tilde{\Phi}_{CA,F}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}-\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\frac{\tilde{W}_{V,F,q}^{i^{T}}\hat{\Psi}_{V,F,q}^{i}}{2}-\tilde{\Phi}_{C,F}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}-\alpha_{V,F,q}^{i}\mathbb{E}\Bigg\{$$

$$\frac{\|\frac{\tilde{W}_{V,F,q}^{i^{T}}\hat{\Psi}_{V,F,q}^{i}}{2}+\varepsilon_{HJB_{F,q}}^{i}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\Bigg\}++\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{CA,F}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}+\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\tilde{\Phi}_{C,F}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}+$$

$$\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|W_{V,F,q}^{i^{T}}\hat{\Psi}_{V,F,q}^{i}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}+\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\varepsilon_{HJB_{F,q}}^{i}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\} \tag{A15}$$

By dropping several negative terms, the derivation can be simplified to

$$\dot{L}_{V,F,q}^{i}(t) \leq -\frac{3}{4}\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^{i}\|^{2}\|\tilde{W}_{V,F,q}^{i}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}+B_{W,V,F,q}^{i}(t) \tag{A16}$$

where

$$B_{W,V,F,q}^{i}(t)=\begin{bmatrix}\alpha_{V,F,q}^{i}[l_{\phi_{CA,F}^{i}}+l_{\Phi_{C,F}^{i}}+l_{\Psi_{V,F,q}^{i}}\mathbb{E}\{\|W_{V,F,q}^{i}\|^{2}\}\|\tilde{\rho}_{F,q}^{i}\|^{2}][1+\mathbb{E}\\ \{\|\hat{\Psi}_{V,F,q}^{i}\|^{2}\}]^{-1}+\alpha_{V,F,q}^{i}\mathbb{E}\left\{\frac{\|\varepsilon_{HJB_{F,q}}^{i}\|^{2}}{1+\|\hat{\Psi}_{V,F,q}^{i}\|^{2}}\right\}\end{bmatrix} \tag{A17}$$

Here, $l_{\Phi_{CA,F}^{i}}$, $l_{\Phi_{C,F}^{i}}$, and $l_{\Psi_{V,F,q}^{i}}$ are the Lipschitz constants, and $\tilde{\rho}_{F,q}^{i}$ is the mass estimation bound.

The follower–critic NN weight estimation error will be UUB, and the bound is

$$\mathbb{E}\left\{\|\tilde{W}_{V,F,q}^{i}\|\right\} \leq \frac{2}{\sqrt{3}}\mathbb{E}\left\{\sqrt{\frac{(1+\|\hat{\Psi}_{V,p,i}\|^{2})}{\alpha_{V,p,i}\|\hat{\Psi}_{V,i}\|^{2}}B_{W,V,F,q}^{i}}\right\} \equiv b_{W,V,F,q}^{i} \tag{A18}$$

This completes the proof.

Appendix B. Proof of Theorem 2

Consider the Lyapunov function candidate

$$L_{\rho,F,q}^{i}(t) = \frac{1}{2}\text{tr}\left\{\mathbb{E}\{\tilde{W}_{\rho,F,q}^{i^{T}}\tilde{W}_{\rho,F,q}^{i}\}\right\} \tag{A19}$$

Additionally, the first derivative of the mass NN estimated weight from (48) can be obtained as follows:

$$\mathbb{E}\{\dot{\tilde{W}}_{\rho,F,q}^{i}\} = \mathbb{E}\left\{-\dot{\hat{W}}_{\rho,F,q}^{i}\right\} = \mathbb{E}\left\{\alpha_{\rho,F,q}^{i}\frac{\hat{\Psi}_{\rho,F,q}^{i}(x_{F,q}^{i},\hat{V}_{F,q}^{i},t)e_{FPK_{F,q}}^{i^{T}}}{1+\|\hat{\Psi}_{\rho,F,q}^{i}(x_{F,q}^{i},\hat{V}_{F,q}^{i},t)\|^{2}}\right\} \tag{A20}$$

According to Lyapunov stability analysis, we take the first derivative of Equation (A19) and substitute in the mass NN weight estimation error dynamic from Equation (A20):

$$\dot{L}_{\rho,F,q}^{i}(t) = \text{tr}\left(\mathbb{E}\left\{\tilde{W}_{\rho,F,q}^{i^{T}}\dot{\tilde{W}}_{\rho,F,q}^{i}\right\}\right)\alpha_{\rho,i}^{i}\text{tr}\left(\mathbb{E}\left\{\tilde{W}_{\rho,i}^{T}\frac{\hat{\Psi}_{\rho,F,q}^{i}(x_{F,q}^{i},\hat{V}_{F,q}^{i},t)e_{FPK_{F,q}}^{i^{T}}}{1+\|\hat{\Psi}_{\rho,F,q}^{i}(x_{F,q}^{i},\hat{V}_{F,q}^{i},t)\|^{2}}\right\}\right) \tag{A21}$$

Let

$$\hat{\Psi}_{\rho,F,q}^{i} = \Psi_{\rho,F,q}^{i}(x_{F,q}^{i},\hat{V}_{F,q}^{i},t) \ ; \ \tilde{\Psi}_{\rho,F,q}^{i} = \Psi_{\rho,F,q}^{i}(x_{F,q}^{i},\tilde{V}_{F,q}^{i},t)$$

Now, by substituting Equation (44) into Equation (A21), we can obtain

$$\dot{L}^i_{\rho,F,q}(t) \le \alpha^i_{\rho,F,q}\text{tr}\left(\mathbb{E}\left\{\left\{\frac{\tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\left\{-\tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q} - W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q} - \varepsilon^i_{\text{FPK}_{F,q}}\right\}\right\}\right)$$

$$\le -\alpha^i_{\rho,F,q}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}\frac{\hat{\Psi}^i_{\rho,F,q}\hat{\Psi}^{i^T}_{\rho,F,q}\tilde{W}^i_{\rho,F,q}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\right) - \alpha^i_{\rho,F,q}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}\frac{\hat{\Psi}^i_{\rho,F,q}\check{\Psi}^{i^T}_{\rho,F,q}W^i_{\rho,F,q}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\right)$$

$$-\alpha^i_{\rho,F,q}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}\frac{\hat{\Psi}^i_{\rho,F,q}\varepsilon^{i^T}_{\text{FPK}_{F,q}}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\right) \tag{A22}$$

Now, the triangle inequality properties are applied to Equation (A22):

$$\dot{L}^i_{\rho,F,q}(t) \le -\frac{1}{2}\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \frac{1}{4}\alpha^i_{\rho,F,q}\times\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \alpha^i_{\rho,F,q}\text{tr}$$

$$\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}\times\frac{\hat{\Psi}^i_{\rho,F,q}\check{\Psi}^{i^T}_{\rho,F,q}W^i_{\rho,F,q}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\right) - \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \frac{1}{4}\alpha^i_{\rho,F,q}\mathbb{E}\Big\{$$

$$\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\Big\} - \alpha^i_{\rho,F,q}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}\frac{\hat{\Psi}^i_{\rho,F,q}\varepsilon^{i^T}_{\text{FPK}_{F,q}}}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\right) - \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\varepsilon^i_{\text{FPK}_{F,q}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}$$

$$+\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} + \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\varepsilon^i_{\text{FPK}_{F,q}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} \tag{A23}$$

Next, Equation (A23) can be simplified as follows:

$$\dot{L}^i_{\rho,F,q}(t) \le -\frac{1}{2}\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \alpha^i_{\rho,F,q}\mathbb{E}\Big\{\frac{1}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\Big[\frac{1}{4}\|\hat{\Psi}^i_{\rho,F,q}\|^2$$

$$\|\tilde{W}^i_{\rho,F,q}\|^2 + \text{tr}\{\tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q}\check{\Psi}^{i^T}_{\rho,F,q}W^i_{\rho,F,q}\} + \|W^T_{\rho,i}\check{\Psi}_{\rho,i}\|^2\Big]\Big\} - \alpha^i_{\rho,F,q}\mathbb{E}\Big\{\frac{1}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}$$

$$\times\Big[\frac{1}{4}\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2 + \tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q}\varepsilon^{i^T}_{\text{FPK}_{F,q}} + \|\varepsilon^i_{\text{FPK}_{F,q}}\|^2\Big]\Big\} + \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}$$

$$+\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\varepsilon^i_{\text{FPK}_{F,q}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}$$

$$\le -\frac{1}{2}\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\frac{\tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q}}{2} + W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} - \alpha^i_{\rho,F,q}$$

$$\mathbb{E}\left\{\frac{\|\frac{\tilde{W}^{i^T}_{\rho,F,q}\hat{\Psi}^i_{\rho,F,q}}{2} + \varepsilon^i_{\text{FPK}_{F,q}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} + \alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|W^{i^T}_{\rho,F,q}\check{\Psi}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\}\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\varepsilon^i_{\text{FPK}_{F,q}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} \tag{A24}$$

After simplification, Equation (A24) can be written as follows:

$$\dot{L}^i_{\rho,F,q}(t) \le -\frac{1}{2}\alpha^i_{\rho,F,q}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} + B^i_{W,\rho,F,q}(t) \tag{A25}$$

where

$$B^i_{W,\rho,F,q} = \left[\alpha^i_{\rho,F,q} \frac{[l_{\Psi^i_{\rho,F,q}} \mathbb{E}\{\|W^i_{\rho,F,q}\|^2\}]\mathbb{E}\{\|\tilde{V}^i_{F,q}\|^2\}}{1+\mathbb{E}\{\|\hat{\Psi}^i_{\rho,F,q}\|^2\}} + \alpha^i_{\rho,F,q}\mathbb{E}\left\{ \frac{\|\varepsilon^i_{\mathrm{FPK_{F,q}}}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2} \right\} \right] \tag{A26}$$

Here, $l_{\Psi^i_{\rho,F,q}}$ is the Lipschitz constant, and $\tilde{V}^i_{F,q}$ is the critic estimation error bound. The mass NN weight estimation error will be UUB, and the bound is

$$\mathbb{E}\left\{ \|\tilde{W}^i_{\rho,F,q}\| \right\} \leq \sqrt{2}\mathbb{E}\left\{ \sqrt{\frac{(1+\|\hat{\Psi}^i_{\rho,F,q}\|^2)}{\alpha^i_{\rho,F,q}\|\hat{\Psi}^i_{\rho,F,q}\|^2} B^i_{W,\rho,F,q}} \right\} \equiv b^i_{W,\rho,F,q} \tag{A27}$$

This completes the proof.

Appendix C. Proof of Theorem 3

Consider the Lyapunov function candidate for the leader

$$L^i_{u,L}(t) = \frac{1}{2}\mathrm{tr}\left\{ \mathbb{E}\{\tilde{W}^T_{u,i}\tilde{W}_{u,i}\} \right\} \tag{A28}$$

The first derivative of the leader–actor NN estimated weight from (34) can be obtained:

$$\mathbb{E}\{\dot{\tilde{W}}^i_{u,L}\} = \mathbb{E}\left\{ -\dot{\hat{W}}^i_{u,L} \right\} = \mathbb{E}\left\{ \alpha^i_{u,L} \frac{\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)e^{iT}_{\mathrm{u,i}}}{1+\|\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\|^2} \right\} \tag{A29}$$

According to the Lyapunov stability analysis, we take the first derivative of Equation (A28) and substitute Equation (A29):

$$\dot{L}^i_{u,L}(t) = \mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L}\dot{\tilde{W}}^i_{u,L} \right\} \right) = \alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)e^{iT}_{\mathrm{u,i}}}{1+\|\hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i)\|^2} \right\} \right) \tag{A30}$$

Let

$$\hat{\phi}^i_{u,L} = \hat{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) \; ; \; \tilde{\phi}^i_{u,L} = \tilde{\phi}^i_{u,L}(x^i_L, x^{-i}_L, \bar{\rho}^i) \; ; \; \tilde{V}^i_L = \tilde{V}_i(x^i_L, x^{-i}_L, \bar{\rho}^i)$$

Now, by substituting Equation (32) into Equation (A30), the following equation is obtained:

$$\dot{L}^i_{u,L}(t) \leq \alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \left\{ \frac{\tilde{W}^{iT}_{u,L}\hat{\phi}^i_{u,L}}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} \left\{ -\tilde{W}^{iT}_{u,L}\hat{\phi}^i_{u,L} - W^{iT}_{u,L}\tilde{\phi}^i_{u,L} - \frac{1}{2}R^{i-1}_L G^T_a(x^i_L)\partial_{x^i_L}\tilde{V}^i_L - \varepsilon^i_{\mathrm{u,L}} \right. \right.$$

$$\left. \left. \right\}^T \right\} \right) \leq -\alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}\hat{\phi}^{iT}_{u,L}\tilde{W}^i_{u,L}}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} \right) - \alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}\tilde{\phi}^{iT}_{u,L}W^i_{u,L}}{1+\|\hat{\Psi}^i_{u,L}\|^2} \right\} \right) - \alpha^i_{u,L}\mathrm{tr}$$

$$\times \left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}[\frac{1}{2}R^{i-1}_L G^T_a(x^i_L)\partial_{x^i_L}\tilde{V}^i_L]}{1+\|\hat{\phi}_{u,i}\|^2} \right\} \right) - \alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}\varepsilon^{iT}_{\mathrm{u,L}}}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} \right) \tag{A31}$$

After applying the triangle inequality property, Equation (A31) can be written as follows:

$$\dot{L}^i_{u,L}(t) \leq -\frac{1}{4}\mathbb{E}\left\{ \alpha^i_{u,L} \frac{\|\hat{\phi}^i_{u,L}\|^2\|\tilde{W}^i_{u,L}\|^2}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} - \frac{1}{4}\mathbb{E}\left\{ \alpha^i_{u,L} \frac{\|\hat{\phi}^i_{u,L}\|^2\|\tilde{W}^i_{u,L}\|^2}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} - \alpha^i_{u,L}\mathrm{tr}\left(\mathbb{E}\left\{ \right. \right.$$

$$\left. \left. \tilde{W}^{iT}_{u,L} \frac{\hat{\phi}^i_{u,L}\tilde{\phi}^{iT}_{u,L}W^i_{u,L}}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} \right) - \alpha^i_{u,L}\mathbb{E}\left\{ \frac{\|W^{iT}_{u,L}\tilde{\phi}^i_{u,L}\|^2}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} + \alpha^i_{u,L}\mathbb{E}\left\{ \frac{\|W^{iT}_{u,L}\tilde{\phi}^i_{u,L}\|^2}{1+\|\hat{\phi}^i_{u,L}\|^2} \right\} - \frac{1}{4}\alpha^i_{u,L}$$

$$\mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \alpha_{u,L}^i \mathrm{tr}\Big(\mathbb{E}\Big\{ \tilde{W}_{u,L}^{i^T} \frac{\hat{\phi}_{u,L}^i [\frac{1}{2} R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i]}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}\Big) - \alpha_{u,L}^i \mathbb{E}$$

$$\times \Big\{ \frac{\|\frac{1}{2} R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\frac{1}{2} R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \frac{1}{4} \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}$$

$$- \alpha_{u,L}^i \mathrm{tr}\Big(\mathbb{E}\Big\{ \tilde{W}_{u,L}^{i^T} \frac{\hat{\phi}_{u,L}^i \varepsilon_{\mathbf{u,L}}^{i^T}}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}\Big) - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}$$

$$\leq -\frac{1}{4} \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{1}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big[\frac{1}{4} \times \|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2 + \mathrm{tr}\{ \tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i}$$

$$\tilde{\phi}_{u,L}^{i^T} W_{u,L}^i\} + \|W_{u,L}^{i^T} \tilde{\phi}_{u,L}^i\|^2 \Big] \Big\} - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{1}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big[\frac{1}{4} \times \|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2 + \mathrm{tr}\{ \tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i}$$

$$\Big[\frac{1}{2} R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i] \} + \|\frac{1}{2} R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2 \Big] \Big\} - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{1}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big[\frac{1}{4} \|\hat{\phi}_{u,L}^i\|^2$$

$$\|\tilde{W}_{u,L}^i\|^2 + \tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i \varepsilon_{\mathbf{u,L}}^{i^T} + \|\varepsilon_{\mathbf{u,i}}^i\|^2 \Big] \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|W_{u,L}^{i^T} \tilde{\phi}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{$$

$$\frac{\|R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{4(1 + \|\hat{\phi}_{u,L}^i\|^2)} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}$$

$$\leq -\frac{1}{4} \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\frac{\tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i}{2} + W_{u,L}^{i^T} \tilde{\phi}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \alpha_{u,L}^i \mathbb{E}$$

$$\Big\{ \frac{\|\frac{\tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i}{2} + R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} - \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\frac{\tilde{W}_{u,L}^{i^T} \hat{\phi}_{u,L}^i}{2} + \varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|W_{u,L}^{i^T} \tilde{\phi}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\}$$

$$+ \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{(1 + \|\hat{\phi}_{u,L}^i\|^2)} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} \tag{A32}$$

Dropping several negative terms and simplifying Equation (A32) yields

$$\dot{L}_{u,L}^i(t) \leq -\frac{1}{4} \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|W_{u,L}^{i^T} \tilde{\phi}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + \alpha_{u,i} \mathbb{E}\Big\{$$

$$\frac{\|R_L^{i^{-1}} G_a^T(x_L^i) \partial_{x_L^i} \tilde{V}_L^i\|^2}{(1 + \|\hat{\phi}_{u,L}^i\|^2)} \Big\} + \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\varepsilon_{\mathbf{u,L}}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} \leq -\frac{1}{4} \alpha_{u,L}^i \mathbb{E}\Big\{ \frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\} + B_{W,u,L}^i \tag{A33}$$

where

$$B_{W,u,L}^i(t) = \begin{bmatrix} \frac{\alpha_{u,L}^i}{1 + \|\hat{\phi}_{u,L}^i\|^2} \Big\{ l_{\phi_{u,L}^i} \mathbb{E}\{\|W_{u,L}^i\|^2 \|\bar{\rho}^i\|^2\} + \mathbb{E}\{\|R_L^{i^{-1}} \\ G_a^T(x_L^i)\|^2 \|\tilde{V}_L^i\|^2\} \Big\} + \alpha_{u,L}^i \frac{\mathbb{E}\{\|\varepsilon_{u,L}^i\|^2\}}{1 + \mathbb{E}\{\|\hat{\phi}_{u,L}^i\|^2\}} \end{bmatrix}$$

Here, $\tilde{V}_i$ is the critic estimation error bound. The actor NN weight estimation error will be UUB, and the bound is

$$\mathbb{E}\Big\{ \|\tilde{W}_{u,L}^i\| \Big\} \leq 2\mathbb{E}\Big\{ \sqrt{\frac{(1 + \|\hat{\phi}_{u,L}^i\|^2)}{\alpha_{u,L}^i \|\hat{\phi}_{u,L}^i\|^2} B_{W,u,L}^i} \Big\} \equiv b_{W,u,L}^i \tag{A34}$$

Similarly, we can derive the bound for the follower–actor NN:

$$\dot{L}^i_{u,F,q}(t) \leq -\frac{1}{4}\alpha^i_{u,F,q}\mathbb{E}\left\{\frac{\|\hat{\phi}^i_{u,F,q}\|^2\|\tilde{W}^i_{u,F,q}\|^2}{1+\|\hat{\phi}^i_{u,F,q}\|^2}\right\} + B^i_{W,u,F,q} \tag{A35}$$

where

$$B^i_{W,u,F,q}(t) = \left[\begin{array}{c} \frac{\alpha^i_{u,F,q}}{1+\mathbb{E}\{\|\hat{\phi}^i_{u,F,q}\|^2\}}\left\{l_{\phi^i_{u,F,q}}\mathbb{E}\{\|W^i_{u,F,q}\|^2\|\tilde{\rho}^i_{F,q}\|^2\} + \mathbb{E}\{\|R^{i^{-1}}_{F,q}\right. \\ \left. G^T_s(x^i_{F,q})\|^2 + \|\tilde{V}^i_{F,q}\|^2\}\right\}\alpha^i_{u,F,q}\frac{\mathbb{E}\{\|\varepsilon^i_{u,F,q}\|^2\}}{1+\mathbb{E}\{\|\hat{\phi}^i_{u,F,q}\|^2\}} \end{array}\right]$$

Here, $l^i_{\phi_{u,F,q}}$ represents the Lipschitz constants, $\tilde{\rho}^i_{F,q}$ is the mass estimation bound, and $\tilde{V}^i_{F,q}$ is the critic estimation error bound. The follower–actor NN weight estimation error will be UUB, and the bound is

$$\mathbb{E}\{\|\tilde{W}^i_{u,F,q}\|\} \leq 2\mathbb{E}\left\{\sqrt{\frac{(1+\|\hat{\phi}^i_{u,F,q}\|^2)}{\alpha^i_{u,F,q}\|\hat{\phi}^i_{u,F,q}\|^2}B^i_{W,u,F,q}}\right\} \equiv b^i_{W,u,F,q} \tag{A36}$$

This completes the proof.

Appendix D. Proof of Theorem 4

Consider the Lyapunov function to be

$$L^i_{\text{sys}}(t) = \frac{\beta_1}{2}\text{tr}\left(\mathbb{E}\left\{e^{i^T}_L(t)e^i_L(t)\right\}\right) + \frac{\beta_2}{2}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{V,L}(t)\tilde{W}^i_{V,L}(t)\right\}\right) + \frac{\beta_3}{2}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{u,L}(t)\right.\right.$$
$$\left.\tilde{W}^i_{u,L}(t)\right\}\right) + \frac{\beta_4}{2}\text{tr}\left(\mathbb{E}\left\{e^{i^T}_{F,q}(t)e^i_{F,q}(t)\right\}\right) + \frac{\beta_5}{2}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{V,F,q}(t)\tilde{W}^i_{V,F,q}(t)\right\}\right) + \frac{\beta_6}{2}\text{tr}\left(\mathbb{E}\left\{\right.\right.$$
$$\left.\left.\tilde{W}^{i^T}_{\rho,F,q}(t)\times\tilde{W}^i_{\rho,F,q}(t)\right\}\right) + \frac{\beta_7}{2}\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{u,F,q}(t)\tilde{W}^i_{u,F,q}(t)\right\}\right) \tag{A37}$$

By taking the first derivative and substituting Lemma 1 and Theorems 1–3 given in Equations (A7), (A16), (A25), (A33), and (A35), we have

$$\dot{L}^i_{\text{sys}}(t) = \beta_1\text{tr}\left(\mathbb{E}\left\{e^{i^T}_L(t)\dot{e}^i_L(t)\right\}\right) + \beta_2\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{V,L}(t)\dot{\tilde{W}}^i_{V,L}(t)\right\}\right) + \beta_3\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{u,L}(t)\right.\right.$$
$$\left.\dot{\tilde{W}}^i_{u,L}(t)\right\}\right) + \beta_4\text{tr}\left(\mathbb{E}\left\{e^{i^T}_{F,q}(t)\dot{e}^i_{F,q}(t)\right\}\right) + \beta_5\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{V,F,q}(t)\dot{\tilde{W}}^i_{V,F,q}(t)\right\}\right) + \beta_6\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{\rho,F,q}(t)\right.\right.$$
$$\left.\left.\times\dot{\tilde{W}}^i_{\rho,F,q}(t)\right\}\right) + \beta_7\text{tr}\left(\mathbb{E}\left\{\tilde{W}^{i^T}_{u,F,q}(t)\dot{\tilde{W}}^i_{u,F,q}(t)\right\}\right)$$

$$\leq \beta_1\text{tr}\left(\mathbb{E}\left\{e^{i^T}_L\left[F_{ar}(e^i_L(t)) + G_{ar}(e^i_L(t))u^i_L(t) + \frac{\sqrt{2\nu^i}d\omega^i_L}{dt}\right]\right\}\right) - \beta_1\text{tr}\left(\mathbb{E}\left\{e^{i^T}_L G_{ar}(e^i_L)\tilde{u}^i_L\right\}\right)$$

$$-\frac{2\beta_1}{\gamma_1}\mathbb{E}\left\{\|G_{ar}(e^i_L)\tilde{u}^i_L\|^2\right\} + \frac{2\beta_1}{\gamma_1}\mathbb{E}\left\{\|G_{ar}(e^i_L)\tilde{u}^i_L\|^2\right\} - \frac{\beta_2\alpha^i_{V,L}}{4}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{V,L}\|^2\|\tilde{W}^i_{V,L}\|^2}{1+\|\hat{\Psi}^i_{V,L}\|^2}\right\} +$$

$$\beta_2 B^i_{W,V,L} - \frac{\beta_3\alpha^i_{u,L}}{4}\mathbb{E}\left\{\frac{\|\hat{\phi}^i_{u,L}\|^2\|\tilde{W}^i_{u,L}\|^2}{1+\|\hat{\phi}^i_{u,L}\|^2}\right\} + \beta_3 B^i_{W,u,L} + \beta_4\text{tr}\left(\mathbb{E}\left\{e^{i^T}_{F,q}\left[F_{sr}(e^i_{F,q}(t)) + G_{sr}(e^i_{F,q}(t))\right.\right.\right.$$

$$\left.\left.\left.u^i_{F,q}(t) + \frac{\sqrt{2\nu^i}d\omega^i_{F,q}}{dt}\right]\right\}\right) - \beta_4\text{tr}\left(\mathbb{E}\left\{e^{i^T}_{F,q}G_{sr}(e^i_{F,q})\tilde{u}^i_{F,q}\right\}\right) - \frac{2\beta_4}{\gamma_2}\mathbb{E}\left\{\|G_{sr}(e^i_{F,q})\tilde{u}^i_{F,q}\|^2\right\} + \frac{2\beta_4}{\gamma_2}$$

$$\mathbb{E}\left\{\|G_{sr}(e^i_{F,q})\tilde{u}^i_{F,q}\|^2\right\} - \frac{\beta_5 3\alpha^i_{V,F,q}}{4}\mathbb{E}\left\{\frac{\|\hat{\Psi}^i_{V,F,q}\|^2\|\tilde{W}^i_{V,F,q}\|^2}{1+\|\hat{\Psi}^i_{V,F,q}\|^2}\right\} + \beta_5 B^i_{W,V,F,q} - \frac{\beta_6\alpha^i_{\rho,F,q}}{2}\mathbb{E}$$

$$\left\{\frac{\|\hat{\Psi}^i_{\rho,F,q}\|^2\|\tilde{W}^i_{\rho,F,q}\|^2}{1+\|\hat{\Psi}^i_{\rho,F,q}\|^2}\right\} + \beta_6 B^i_{W,\rho,F,q} - \frac{\beta_7\alpha^i_{u,F,q}}{4}\mathbb{E}\left\{\frac{\|\hat{\phi}^i_{u,F,q}\|^2\|\tilde{W}^i_{u,F,q}\|^2}{1+\|\hat{\phi}^i_{u,F,q}\|^2}\right\} + \beta_7 B^i_{W,u,F,q}$$

$$\leq -\frac{\gamma_1 \beta_1}{2} \mathbb{E}\left\{\|e_L^i\|^2\right\} + \frac{2\beta_1 g_{l1}^2}{\gamma_1} \mathbb{E}\left\{\|\tilde{u}_L^i\|^2\right\} - \frac{\beta_2 \alpha_{V,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2 \|\tilde{W}_{V,L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\} + \beta_2 B_{W,V,L}^i -$$

$$\frac{\beta_3 \alpha_{u,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2}\right\} + \beta_3 B_{W,u,L}^i - \frac{\gamma_2 \beta_4}{2} \mathbb{E}\left\{\|e_{F,q}^i\|^2\right\} + \frac{2\beta_4}{\gamma_2} g_{l2}^2 \mathbb{E}\left\{\|\tilde{u}_{F,q}^i\|^2\right\}$$

$$- \frac{\beta_5 3\alpha_{V,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2 \|\tilde{W}_{V,F,q}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}^i\|^2}\right\} + \beta_5 B_{W,V,F,q}^i - \frac{\beta_6 \alpha_{\rho,F,q}^i}{2} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{\rho,F,q}^i\|^2 \|\tilde{W}_{\rho,F,q}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\} + \beta_6$$

$$B_{W,\rho,F,q}^i - \frac{\beta_7 \alpha_{u,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,F,q}^i\|^2 \|\tilde{W}_{u,F,q}^i\|^2}{1 + \|\hat{\phi}_{u,F,q}^i\|^2}\right\} + \beta_7 B_{W,u,F,q}^i \tag{A38}$$

where g_{l1} and g_{l2} are the Lipschitz constants of the dynamic equations $G_{ar}(e_L^i)$ and $G_{sr}(e_{F,q}^i)$, respectively. Now, by substituting Equations (49)–(53) into Equation (A38), Equation (A38) can be represented and simplified as follows:

$$\dot{L}_{sys}^i(t) \leq b_1 + b_3 + 3b_4 \mathbb{E}\left\{\|\tilde{W}_{V,L}^{i^T}\|^2 \|\hat{\phi}_{V,L}^i\|^2\right\} + [3b_4 l_{\psi_{V,L}^i}^2 + b_2] \bar{b}_{\rho,F,q}^{i2} + 3b_4 \|\varepsilon_{\text{HJB}_L}^i\|^2 + b_5 +$$

$$b_7 3b_8 \|\tilde{W}_{V,F,q}^{i^T}\|^2 \|\hat{\phi}_{V,F,q}^i\|^2 + 2[3b_8 l_{\psi_{V,F,q}^i}^2 \|W_{V,F,q}^i\| + b_6] \|\tilde{W}_{\rho,F,q}^{i^T}\|^2 \|\psi_{\rho,F,q}^i\|^2 + [3b_8 l_{\psi_{V,F,q}^i}^2$$

$$\|W_{V,F,q}^i\| + b_6] \|\varepsilon_{\text{FPK}_{F,q}}^i\|^2 + 3b_8 \|\varepsilon_{\text{HJB}_{F,q}}^i\|^2 \tag{A39}$$

where

$$b_1 = -\frac{\gamma_1 \beta_1}{2} \mathbb{E}\left\{\|e_L^i\|^2\right\} + \frac{2\beta_1 g_{l1}^2}{\gamma_1} \mathbb{E}\left\{\|\tilde{u}_L^i\|^2\right\} - \frac{\beta_2 \alpha_{V,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2 \|\tilde{W}_{V,L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\}$$

$$b_2 = \beta_2 \alpha_{V,L}^i \left[\frac{[l_{\Phi_{\text{CP,L}}^i} + l_{\Psi_{V,L}^i} \mathbb{E}\{\|W_{V,L}^i\|^2\}]}{1 + \mathbb{E}\{\|\hat{\Psi}_{V,L}^i\|^2\}}\right]$$

$$b_3 = \frac{\beta_3 \alpha_{u,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,L}^i\|^2 \|\tilde{W}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2}\right\} + \beta_2 \alpha_{V,L}^i \mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\} + \beta_3 \alpha_{u,i} \frac{\mathbb{E}\{\|\varepsilon_{u,L}^i\|^2\}}{1 + \mathbb{E}\{\|\hat{\phi}_{u,L}^i\|^2\}}$$

$$b_4 = \frac{\beta_3 \alpha_{u,L}^i}{1 + \|\hat{\phi}_{u,L}^i\|^2} \left\{l_{\psi_{u,L}^i} \mathbb{E}\{\|W_{u,L}^i\|^2 \|\bar{\rho}^i\|^2\} + \mathbb{E}\{\|R_L^{i^{-1}} G_a^T(x_L^i)\|^2\}\right\}$$

$$b_5 = -\frac{\gamma_2 \beta_4}{2} \mathbb{E}\left\{\|e_{F,q}^i\|^2\right\} + \frac{2\beta_4}{\gamma_2} g_{l2}^2 \mathbb{E}\left\{\|\tilde{u}_{F,q}^i\|^2\right\} - \frac{\beta_5 3\alpha_{V,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2 \|\tilde{W}_{V,F,q}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}^i\|^2}\right\}$$

$$b_6 = \beta_5 \alpha_{V,F,q}^i [l_{\phi_{\text{CA,F}}^i} + l_{\Phi_{\text{C,F}}^i} + l_{\Psi_{V,F,q}^i} \mathbb{E}\{\|W_{V,F,q}^i\|^2\}][1 + \mathbb{E}\{\|\hat{\Psi}_{V,F,q}^i\|^2\}]^{-1}$$

$$b_7 = \frac{\beta_6 \alpha_{\rho,F,q}^i}{2} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{\rho,F,q}^i\|^2 \|\tilde{W}_{\rho,F,q}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\} - \frac{\beta_7 \alpha_{u,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,F,q}^i\|^2 \|\tilde{W}_{u,F,q}^i\|^2}{1 + \|\hat{\phi}_{u,F,q}^i\|^2}\right\} + \beta_5 \alpha_{V,F,q}^i \mathbb{E}$$

$$\left\{\frac{\|\varepsilon_{\text{HJB}_{F,q}}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}^i\|^2}\right\} + \beta_6 \alpha_{\rho,F,q}^i \mathbb{E}\left\{\frac{\|\varepsilon_{\text{FPK}_{F,q}}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\} + \beta_7 \alpha_{u,F,q}^i \frac{\mathbb{E}\{\|\varepsilon_{u,F,q}^i\|^2\}}{1 + \mathbb{E}\{\|\hat{\phi}_{u,F,q}^i\|^2\}}$$

$$b_8 = \beta_6 \alpha_{\rho,F,q}^i \frac{[l_{\Psi_{\rho,F,q}^i} \mathbb{E}\{\|W_{\rho,F,q}^i\|^2\}]}{1 + \mathbb{E}\{\|\hat{\Psi}_{\rho,F,q}^i\|^2\}} + \beta_7 \frac{\alpha_{u,F,q}^i}{1 + \mathbb{E}\{\|\hat{\phi}_{u,F,q}^i\|^2\}}$$

$$\left\{l_{\phi_{u,F,q}^i} \mathbb{E}\{\|W_{u,F,q}^i\|^2 \|\tilde{\rho}_{F,q}^i\|^2\} + \mathbb{E}\{\|R_{F,q}^{i^{-1}} G_s^T(x_{F,q}^i)\|^2\}\right\}$$

In addition, let

$$\varepsilon_{\text{NHJB}_L}^i = \alpha_{V,L}^i \mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\}; \ \varepsilon_{\text{N}_{u,L}}^i = \alpha_{u,L}^i \frac{\mathbb{E}\{\|\varepsilon_{u,L}^i\|^2\}}{1 + \mathbb{E}\{\|\hat{\phi}_{u,L}^i\|^2\}}$$

$$\varepsilon_{\text{NHJB}_{F,q}^i} = \alpha_{V,F,q}^i \mathbb{E}\left\{\frac{\|\varepsilon_{\text{HJB}_{F,q}}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}\|^2}\right\} \; ; \; \varepsilon_{\text{N}_{u,F,q}^i} = \alpha_{u,F,q}^i \frac{\mathbb{E}\{\|\varepsilon_{u,F,q}^i\|^2\}}{1 + \mathbb{E}\{\|\hat{\phi}_{u,F,q}^i\|^2\}}$$

$$\varepsilon_{\text{NFPK}_{F,q}^i} = \alpha_{\rho,F,q}^i \mathbb{E}\left\{\frac{\|\varepsilon_{\text{FPK}_{F,q}}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\}$$

Again, Equation (A39) can be written as follows:

$$\dot{L}_{sys}^i(t) \le -\frac{\gamma_1 \beta_1}{2} \mathbb{E}\{\|e_L^i\|^2\} - \left[\frac{\beta_2 \alpha_{V,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\} - 3b_4\right] \mathbb{E}\{\|\tilde{W}_{V,L}^i\|^2\} - \left[\frac{\beta_3 \alpha_{u,L}^i}{4}\right.$$

$$\mathbb{E}\left\{\frac{\|\hat{\phi}_{u,i}\|^2}{1 + \|\hat{\phi}_{u,i}\|^2}\right\} - \frac{6\beta_1 g_{l1}^2}{\gamma_1} \|\hat{\Psi}_{u,L}^i\|^2\Big] \|\tilde{W}_{u,L}^i\|^2 - \frac{\gamma_2 \beta_4}{2} \mathbb{E}\{\|e_{F,q}^i\|^2\} - \left[\frac{3\beta_5 \alpha_{V,F,q}^i}{4}\right.$$

$$\mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2 \|\tilde{W}_{V,F,q}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}^i\|^2}\right\} - 3b_8 \|\hat{\phi}_{V,F,q}^i\|^2\Big] \|\tilde{W}_{V,F,q}^i\|^2 - \left[\frac{\beta_7 \alpha_{u,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,F,q}^i\|^2}{1 + \|\hat{\phi}_{u,F,q}^i\|^2}\right\}\right.$$

$$\left. - \frac{6\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|\hat{\phi}_{u,F,q}^i\|^2\}\right] \mathbb{E}\{\|\tilde{W}_{u,F,q}^i\|^2\} - \left[\frac{\beta_6 \alpha_{\rho,F,q}^i}{2} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{\rho,F,q}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\} - \frac{12\beta_4 g_{l2}^2}{\gamma_2}\right.$$

$$\mathbb{E}\{\|W_{u,F,q}^i\|^2\} l_{\Psi_{u,F,q}^i}^2 \mathbb{E}\{\|\hat{\phi}_{\rho,F,q}^i\|^2\} - 6b_8 l_{\Psi_{V,F,q}^i}^2 \mathbb{E}\{\|W_{V,F,q}^i\| \|\hat{\Psi}_{\rho,F,q}^i\|^2\} - 2b_6 \mathbb{E}\{\|\hat{\Psi}_{\rho,F,q}^i\|^2\}\Big]$$

$$\mathbb{E}\{\|\tilde{W}_{\rho,F,q}^i\|^2\} + \frac{6\beta_1 g_{l1}^2}{\gamma_1} \mathbb{E}\{\|\varepsilon_{u,L}^i\|^2\} + \beta_2 \varepsilon_{\text{NHJB}_L^i} + \beta_3 \varepsilon_{\text{N}_{u,L}^i} + 3b_4 \mathbb{E}\{\|\varepsilon_{\text{HJB}_L}^i\|^2\} \frac{12\beta_4 g_{l2}^2}{\gamma_2}$$

$$\mathbb{E}\{\|W_{u,F,q}^i\|^2\} l_{\phi_{u,F,q}^i}^2 \mathbb{E}\{\|\varepsilon_{\text{FPK}_{F,q}}^i\|^2\} + \frac{6\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|\varepsilon_{u,F,q}^i\|^2\} + \beta_5 \varepsilon_{\text{NHJB}_{F,q}^i} + \beta_6 \varepsilon_{\text{NFPK}_{F,q}^i}$$

$$+ 6b_8 l_{\Psi_{V,F,q}^i}^2 \mathbb{E}\{\|W_{V,F,q}^i\| \|\varepsilon_{\text{FPK}_{F,q}}^i\|^2\} + 2b_6 \mathbb{E}\{\|\varepsilon_{\text{FPK}_{F,q}}^i\|^2\} + 3b_8 \|\varepsilon_{\text{HJB}_L}^i\|^2$$

$$\le -\frac{\gamma_1 \beta_1}{2} \mathbb{E}\{\|e_L^i\|^2\} - \kappa_{V,L}^i \mathbb{E}\{\|\tilde{W}_{V,L}^i\|^2\} - \kappa_{u,L}^i \mathbb{E}\{\|\tilde{W}_{u,L}^i\|^2\} - \frac{\gamma_2 \beta_4}{2} \mathbb{E}\{\|e_{F,q}^i\|^2\} - \kappa_{V,F,q}^i$$

$$\mathbb{E}\{\|\tilde{W}_{V,F,q}^i\|^2\} - \kappa_{u,F,q}^i \mathbb{E}\{\|\tilde{W}_{u,F,q}^i\|^2\} - \kappa_{\rho,F,q}^i \mathbb{E}\{\|\tilde{W}_{\rho,F,q}^i\|^2\} + \varepsilon_{CS} \tag{A40}$$

with

$$\kappa_{V,L}^i = \left[\frac{\beta_2 \alpha_{V,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,L}^i\|^2}{1 + \|\hat{\Psi}_{V,L}^i\|^2}\right\} - 3b_4\right]$$

$$\kappa_{u,L}^i = \frac{\beta_3 \alpha_{u,L}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,L}^i\|^2}{1 + \|\hat{\phi}_{u,L}^i\|^2}\right\} - \frac{6\beta_1 g_{l1}^2}{\gamma_1} \mathbb{E}\{\|\hat{\phi}_{u,L}^i\|^2\}$$

$$\kappa_{V,F,q}^i = \frac{3\beta_5 \alpha_{v,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{V,F,q}^i\|^2 \|\tilde{W}_{V,F,q}^i\|^2}{1 + \|\hat{\Psi}_{V,F,q}^i\|^2}\right\} - 3b_8 \mathbb{E}\{\|\hat{\Psi}_{V,F,q}^i\|^2\}$$

$$\kappa_{u,F,q}^i = \frac{\beta_7 \alpha_{u,F,q}^i}{4} \mathbb{E}\left\{\frac{\|\hat{\phi}_{u,F,q}^i\|^2}{1 + \|\hat{\phi}_{u,F,q}^i\|^2}\right\} - \frac{6\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|\hat{\phi}_{u,F,q}^i\|^2\}$$

$$\kappa_{\rho,F,q}^i = \frac{\beta_6 \alpha_{\rho,F,q}^i}{2} \mathbb{E}\left\{\frac{\|\hat{\Psi}_{\rho,F,q}^i\|^2}{1 + \|\hat{\Psi}_{\rho,F,q}^i\|^2}\right\} - \frac{12\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|W_{u,F,q}^i\|^2\} l_{\phi_{u,F,q}^i}^2 \mathbb{E}\{\|\hat{\Psi}_{\rho,F,q}^i\|^2\} - 6b_8 l_{\Psi_{V,F,q}^i}^2$$

$$\mathbb{E}\{\|W_{V,F,q}^i\| \|\hat{\Psi}_{\rho,F,q}^i\|^2\} - 2b_6 \mathbb{E}\{\|\hat{\Psi}_{\rho,F,q}^i\|^2\}$$

$$\varepsilon_{CS} = \frac{6\beta_1 g_{l1}^2}{\gamma_1} \mathbb{E}\{\|\varepsilon_{u,L}^i\|^2\} + \beta_2 \varepsilon_{\text{NHJB}_L^i} + \beta_3 \varepsilon_{\text{N}_{u,L}^i} + 3b_4 \|\varepsilon_{\text{HJB}_L}^i\|^2 \frac{12\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|W_{u,F,q}^i\|^2\}$$

$$l_{\phi_{u,F,q}^i}^2 \mathbb{E}\{\|\varepsilon_{\text{FPK}_{F,q}}^i\|^2\} + \frac{6\beta_4 g_{l2}^2}{\gamma_2} \mathbb{E}\{\|\varepsilon_{u,F,q}^i\|^2\} + \beta_5 \varepsilon_{\text{NHJB}_{F,q}^i} + \beta_6 \varepsilon_{\text{NFPK}_{F,q}^i} + 6b_8 l_{\Psi_{V,F,q}^i}^2$$

$$\mathbb{E}\big\{\|W_{V,F,q}^i\|\|\varepsilon_{\mathrm{FPK_{F,q}}}^i\|^2\big\} + 2b_6\mathbb{E}\big\{\|\varepsilon_{\mathrm{FPK_{F,q}}}^i\|^2\big\} + 3b_8\mathbb{E}\big\{\|\varepsilon_{\mathrm{HJB_{F,q}}}^i\|^2\big\}$$

The derivation of the Lyapunov function $\dot{L}_{sys}^i(t)$ is less than zero outside a compact set; in other words, we have

$$\mathbb{E}\big\{\|e_L^i\|\big\} > \sqrt{\frac{2}{\gamma_1\beta_1}\varepsilon_{CS}} \;;\; \mathbb{E}\big\{\|\tilde{W}_{V,L}^i\|\big\} > \sqrt{\frac{1}{k_{V,L}^i}\varepsilon_{CS}} \;;\; \mathbb{E}\big\{\|\tilde{W}_{u,L}^i\|\big\} > \sqrt{\frac{1}{k_{u,L}^i}\varepsilon_{CS}}$$

$$\mathbb{E}\big\{\|e_{F,q}^i\|\big\} > \sqrt{\frac{2}{\gamma_2\beta_5}\varepsilon_{CS}} \;;\; \mathbb{E}\big\{\|\tilde{W}_{V,F,q}^i\|\big\} > \sqrt{\frac{1}{k_{V,F,q}^i}\varepsilon_{CS}} \;;\; \mathbb{E}\big\{\|\tilde{W}_{u,F,q}^i\|\big\} > \sqrt{\frac{1}{k_{u,F,q}^i}\varepsilon_{CS}}$$

$$\mathbb{E}\big\{|\tilde{W}_{\rho,F,q}^i\|\big\} > \sqrt{\frac{1}{k_{\rho,F,q}^i}\varepsilon_{CS}}$$

This completes the proof.

References

1. Topaz, C.M.; Bertozzi, A.L. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.* **2004**, *65*, 152–174. [CrossRef]
2. Okubo, A. Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. *Adv. Biophys.* **1986**, *22*, 1–94. [CrossRef] [PubMed]
3. Toner, J.; Tu, Y. Flocks, herds, and schools: A quantitative theory of flocking. *Phys. Rev. E* **1998**, *58*, 4828. [CrossRef]
4. Kube, C.R.; Bonabeau, E. Cooperative transport by ants and robots. *Robot. Auton. Syst.* **2000**, *30*, 85–101. [CrossRef]
5. Li, W.; Shen, W. Swarm behavior control of mobile multi-robots with wireless sensor networks. *J. Netw. Comput. Appl.* **2011**, *34*, 1398–1407. [CrossRef]
6. Cao, L.; Cai, Y.; Yue, Y. Swarm intelligence-based performance optimization for mobile wireless sensor networks: Survey, challenges, and future directions. *IEEE Access* **2019**, *7*, 161524–161553. [CrossRef]
7. Berman, S.; Halász, A.; Hsieh, M.A.; Kumar, V. Optimized stochastic policies for task allocation in swarms of robots. *IEEE Trans. Robot.* **2009**, *25*, 927–937. [CrossRef]
8. Bayındır, L. A review of swarm robotics tasks. *Neurocomputing* **2016**, *172*, 292–321. [CrossRef]
9. Jevtic, A.; Gutiérrez, A.; Andina, D.; Jamshidi, M. Distributed bees algorithm for task allocation in swarm of robots. *IEEE Syst. J.* **2011**, *6*, 296–304. [CrossRef]
10. Engelen, S.; Gill, E.; Verhoeven, C. On the reliability, availability, and throughput of satellite swarms. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 1027–1037. [CrossRef]
11. Xu, D.; Zhang, X.; Zhu, Z.; Chen, C.; Yang, P. Behavior-based formation control of swarm robots. *Math. Probl. Eng.* **2014**, *2014*, 205759. [CrossRef]
12. Soni, A.; Hu, H. Formation control for a fleet of autonomous ground vehicles: A survey. *Robotics* **2018**, *7*, 67. [CrossRef]
13. Tahir, A.; Böling, J.; Haghbayan, M.H.; Toivonen, H.T.; Plosila, J. Swarms of unmanned aerial vehicles—A survey. *J. Ind. Inf. Integr.* **2019**, *16*, 100106. [CrossRef]
14. Zhu, B.; Xie, L.; Han, D. Recent developments in control and optimization of swarm systems: A brief survey. In Proceedings of the 2016 12th IEEE international conference on control and automation (ICCA), Kathmandu, Nepal, 1–3 June 2016; pp. 19–24.
15. Lan, X.; Liu, Y.; Zhao, Z. Cooperative control for swarming systems based on reinforcement learning in unknown dynamic environment. *Neurocomputing* **2020**, *410*, 410–418. [CrossRef]
16. Skobelev, P.; Budaev, D.; Gusev, N.; Voschuk, G. Designing multi-agent swarm of uav for precise agriculture. In Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems, Toledo, Spain, 20–22 June 2018; pp. 47–59.
17. Kada, B.; Khalid, M.; Shaikh, M.S. Distributed cooperative control of autonomous multi-agent UAV systems using smooth control. *J. Syst. Eng. Electron.* **2020**, *31*, 1297–1307. [CrossRef]
18. Xia, Z.; Du, J.; Wang, J.; Jiang, C.; Ren, Y.; Li, G.; Han, Z. Multi-Agent Reinforcement Learning Aided Intelligent UAV Swarm for Target Tracking. *IEEE Trans. Veh. Technol.* **2022**, *71*, 931–945. [CrossRef]
19. Zhao, W.; Chu, H.; Zhang, M.; Sun, T.; Guo, L. Flocking control of fixed-wing UAVs with cooperative obstacle avoidance capability. *IEEE Access* **2019**, *7*, 17798–17808. [CrossRef]
20. Zhou, Z.; Xu, H. A Novel Mean-Field-Game-Type Optimal Control for Very Large-Scale Multiagent Systems. *IEEE Trans. Cybern.* **2022**, *52*, 5197–5208. [CrossRef]

21. Mehlfuhrer, C.; Caban, S.; Rupp, M. Cellular system physical layer throughput: How far off are we from the Shannon bound? *IEEE Wirel. Commun.* **2011**, *18*, 54–63. [CrossRef]

22. Branzei, R.; Dimitrov, D.; Tijs, S. *Models in Cooperative Game Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008; Volume 556.

23. Gulzar, M.M.; Rizvi, S.T.H.; Javed, M.Y.; Munir, U.; Asif, H. Multi-agent cooperative control consensus: A comparative review. *Electronics* **2018**, *7*, 22. [CrossRef]

24. Zhou, Z.; Xu, H. Decentralized Adaptive Optimal Tracking Control for Massive Autonomous Vehicle Systems With Heterogeneous Dynamics: A Stackelberg Game. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 5654–5663. [CrossRef]

25. Yu, M.; Hong, S.H. A Real-Time Demand-Response Algorithm for Smart Grids: A Stackelberg Game Approach. *IEEE Trans. Smart Grid* **2016**, *7*, 879–888. [CrossRef]

26. Cardaliaguet, P.; Porretta, A. An introduction to mean field game theory. In *Mean Field Games*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 1–158.

27. Yang, Y.; Luo, R.; Li, M.; Zhou, M.; Zhang, W.; Wang, J. Mean field multi-agent reinforcement learning. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 5571–5580.

28. Shiri, H.; Park, J.; Bennis, M. Massive autonomous UAV path planning: A neural network based mean-field game theoretic approach. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; pp. 1–6.

29. Bogachev, V.I.; Krylov, N.V.; Röckner, M.; Shaposhnikov, S.V. *Fokker–Planck–Kolmogorov Equations*; American Mathematical Society: Providence, RI, USA, 2022; Volume 207.

30. Peng, S. Stochastic hamilton–jacobi–bellman equations. *SIAM J. Control Optim.* **1992**, *30*, 284–304. [CrossRef]

31. Murray, J.J.; Cox, C.J.; Lendaris, G.G.; Saeks, R. Adaptive dynamic programming. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2002**, *32*, 140–153. [CrossRef]

32. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [CrossRef]

33. Ju, C.; Son, H.I. Multiple UAV systems for agricultural applications: Control, implementation, and evaluation. *Electronics* **2018**, *7*, 162. [CrossRef]

34. Gupta, J.K.; Egorov, M.; Kochenderfer, M. Cooperative multi-agent control using deep reinforcement learning. In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, Sao Paulo, Brazil, 8–12 May 2017; pp. 66–83.

35. Oroojlooy, A.; Hajinezhad, D. A review of cooperative multi-agent deep reinforcement learning. *Appl. Intell.* **2022**, 1–46. [CrossRef]

36. Zhang, S.; Li, T.; Cheng, X.; Li, J.; Xue, B. Multi-Group Formation Tracking Control for Second-Order Nonlinear Multi-Agent Systems Using Adaptive Neural Networks. *IEEE Access* **2021**, *9*, 168207–168215. [CrossRef]

37. Wu, Z.; Liu, X.; Sun, J.; Wang, X. Multi-group formation tracking control via impulsive strategy. *Neurocomputing* **2020**, *411*, 487–497. [CrossRef]

38. Luo, L.; Wang, X.; Ma, J.; Ong, Y.S. Grpavoid: Multigroup collision-avoidance control and optimization for UAV swarm. *IEEE Trans. Cybern.* **2021**. [CrossRef]

39. Lewis, F.L.; Vrabie, D.; Vamvoudakis, K.G. Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers. *IEEE Control. Syst. Mag.* **2012**, *32*, 76–105.