MDPI

Article

A Cooperative Decision-Making and Control Algorithm for UAV Formation Based on Non-Cooperative Game Theory

Yongkang Jiao [1,2], Wenxing Fu [1], Xinying Cao [2], Kunhu Kou [2], Ji Tang [2], Rusong Shen [2], Yiyang Zhang [3] and Haibo Du [3,*]

[1] Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an 710072, China; jiaoykang@163.com (Y.J.); wenxingfu@nwpu.edu.cn (W.F.)
[2] Aeronautical Operation Institute, Naval Aviation University, Yantai 264001, China; caoxinying1122@sina.com (X.C.); koukunhu@163.com (K.K.); mantis08082024@163.com (J.T.); naau_srs@163.com (R.S.)
[3] School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China; yiyang.zhang@mail.hfut.edu.cn
* Correspondence: haibo.du@hfut.edu.cn; Tel.: +86-1815-606-2198

Abstract: The formation control problem of distributed fixed-wing Unmanned Aerial Vehicles (UAVs) is investigated in this paper. By utilizing the theoretical foundations of non-cooperative game theory, a novel control strategy is introduced, which allows UAVs to autonomously determine the optimal flight trajectory without relying on centralized coordination while concurrently mitigating conflicts with other UAVs. By transforming the UAV model into a double integrator form, the control complexity is reduced. Additionally, the incorporation of a homogeneous differential disturbance observer enhances the UAV's resilience against disturbances during the control process. Through the development and validation of a Nash equilibrium-based algorithm, it is demonstrated that UAVs can sustain a predefined formation flight and autonomously adapt their trajectories in complex environments. Simulations are presented to confirm the efficiency of the proposed method.

Keywords: non-cooperative game theory; distributed control; disturbance observer; fixed-wing unmanned aerial vehicles; formation control; Nash equilibrium; backstepping

1. Introduction

Recently, the utilization of Unmanned Aerial Vehicles (UAVs) has become increasingly prevalent. Specifically, fixed-wing UAVs are prized for their extended endurance and superior speed. Formation flight serves as a pivotal technology for enhancing the efficiency of UAV missions, allowing multiple UAVs to collaborate in tasks such as search, surveillance, mapping, and target localization [1–7]. However, with the escalating complexity of mission environments and the proliferating number of UAVs, traditional centralized control approaches encounter obstacles, including constraints in communication bandwidth, risks of single-point failures, and limitations in real-time processing [8–10]. In contrast, the distributed formation control method rooted in graph theory exhibits strong stability and robustness [11–15].

Unlike cooperative games, non-cooperative participants optimize independently, which is suitable for scenarios where UAVs cannot or will not share information. Assuming each UAV minimizes its cost function based on its own and others' actions, an appropriate game-theoretic model and Nash equilibrium ensure stable equilibrium without external enforcement [16,17]. The communication and computational costs of Nash equilibrium-seeking strategies in non-cooperative games were reduced by introducing an interference graph to describe the interactions between agents within each coalition and by designing Nash equilibrium-seeking strategies that consider the interference graph [18]. A projection algorithm based on continuous-time distributed gradient was proposed to address the

problem where both the cost function and the feasible action set depend on the actions of all participants, and where each participant can only access the information of its neighbors. This algorithm utilizes a leader–follower consensus algorithm to estimate the behavior of each participant [19].

The Disturbance Observer was proposed by Japanese scholar K. Ohnishi in 1987 [20]. It is primarily used to handle external disturbances and uncertainties arising from changes in model parameters in control systems. By observing and compensating for these disturbances, it achieves optimization and stabilization of system performance. A time-varying dynamic modeling method for space double flexible telescopic manipulators, coupled with a fuzzy PI real-time control strategy integrated with a fractional order disturbance observer, was proposed to address the challenges posed by their nonlinear and time-varying characteristics, aiming to enhance operational stability and precision [21]. A finite-time control method is proposed for unmanned ground vehicle trajectory tracking, considering uncertainty and disturbance compensation. Employing a Frenet-coordinate system and a novel coordinate transformation, a finite-time disturbance observer and controller based on an integral-type Lyapunov function are designed [22]. An improved discrete super-twisting control algorithm is proposed, leveraging homogeneous system theory to precisely estimate and compensate for system disturbances, enhancing robustness. An adjustable fractional power parameter is introduced to address sampling constraints, ensuring higher control accuracy in closed-loop systems [23].

In the context of multi-UAV formation control, backstepping can be used to develop controllers that ensure the formation shape and spacing between the UAVs are maintained during flight [24,25]. The effectiveness of backstepping for this complex and challenging control problem has been demonstrated [26]. An adaptive fuzzy tracking control strategy for pure feedback nonlinear multi-agent systems, developed through backstepping, has been proposed [27]. An approach combining backstepping technology with neural networks, introducing a radial basis function neural network with a gradient descent algorithm, has been proposed to address the unknown complex nonlinear problems of UAV systems caused by external disturbances [28]. A distributed formation controller is designed using the backstepping method based on the asymmetric barrier Lyapunov function [29].

This paper introduces a distributed formation control strategy for fixed-wing UAVs based on the non-cooperative game theory. Compared to the aforementioned literature, this paper makes the following innovations. A double integrator model was used to reduce the design complexity of the controller [30]. We innovatively combine the principles of non-cooperative game theory with the backstepping control method to design a distributed formation controller, enabling the states of the multi-UAV system to converge to a Nash equilibrium point. Additionally, a homogeneous differential disturbance observer is employed to observe and mitigate the disturbances generated in the system. The effectiveness of this method has been confirmed by simulations.

2. Preliminaries

2.1. Graph Theory

In a multi-agent communication topology, the graph $G = \{V, E\}$ embodies the connectivity among various agents. Here, $V = \{V_1, V_2, \ldots, V_N\}$ denotes the set of nodes, representing individual agents, and $E = \{E_1, E_2, \ldots, E_M\}$ signifies the set of edges, denoting communication links between these nodes. The presence of an edge from node i to node j is symbolically indicated by the pair $(i, j) \in E$. Furthermore, if there exists a sequence of edges $(i, j_1), (j_1, j_2), \ldots, (j_m, j) \in E$, this implies that there is a directed path enabling communication from node i to node j.

A directed graph G qualifies as strongly connected if, regardless of the pair of nodes chosen, a directed path facilitating communication between them can be identified. The adjacency matrix associated with this directed graph is denoted as A, where each element a_{ij} corresponds to the entry in the i-th row and j-th column. Specifically, $a_{ij} = 1$ when there

exists an edge from node i to node j, i.e., $(i, j) \in E$, while $a_{ij} = 0$ in all other cases. It is important to note that the graph G explicitly disallows self-loops, meaning that $a_{ii} = 0$ for all nodes.

The in-degree matrix, represented as D, is constructed as a diagonal matrix with the degrees of the nodes along its diagonal. Mathematically, it is expressed as $D = \text{diag}\{\deg(V_1), \deg(V_2), \ldots, \deg(V_N)\}$, where the degree $\deg(V_j)$ of a node V_j is calculated as the sum of the entries in the j-th column of the adjacency matrix A, or $\deg(V_j) = \sum_{i=1}^{N} a_{ij}$.

The Laplacian matrix L of the graph G is then defined as the difference between the in-degree matrix D and the adjacency matrix A, formulated as $L = D - A$. In the context of a strongly connected graph $G = \{V, E\}$, the eigenvalue spectrum of the Laplacian matrix L exhibits a unique characteristic: the value 0 emerges as the sole eigenvalue, accompanied by the vector 1_N as the corresponding eigenvector [31].

2.2. Non-Cooperative Game

Defining V as the set of N players, $o_i \in \mathbb{R}^n$ represents the operation of the i-th player, $o = [o_1^T, \ldots, o_N^T]^T$ is the joint action vector, and $f_i(o)$ the cost function associated with the i-th player. In a non-cooperative game, the cost function for each player is determined by the joint actions of all players involved in the game. Each player aims to minimize their own individual cost function, taking into account the actions of the other players. This results in a strategic interaction where the optimal strategy for each player depends on the strategies chosen by the others.

If there exists a joint action vector $o^* = [o_1^{*T}, \ldots, o_N^{*T}]^T$, which means a collection of strategies chosen by each of the N players, such that for any player i within the set $\{1, \ldots, N\}$ and for any alternative strategy o_i that this player might consider, the following condition holds:

$$f_i(o_i, o_{-i}^*) \geq f_i(o_i^*, o_{-i}^*) \tag{1}$$

where o^* is referred to as a Nash equilibrium solution of the non-cooperative game [32]. Here, $o_{-i} = [o_1^T, \ldots, o_{i-1}^T, o_{i+1}^T, \ldots, o_N^T]^T$ represents the actions of all players except the i-th player, highlighting the interdependence of players' decisions in the game.

The aforementioned inequality signifies a crucial property of the Nash equilibrium: it ensures that once the system reaches this state, no individual player has an incentive to deviate from their current strategy, as doing so would not lead to a reduction in their cost function, or equivalently, would not improve their payoff. This principle underlines the stability of the Nash equilibrium as a solution concept in non-cooperative games.

For an unconstrained non-cooperative game among players, the Nash equilibrium is mathematically characterized by the first-order condition $\frac{\partial f_i(o^*)}{\partial o_i} = 0$. This condition implies that at the equilibrium point, the marginal change in the cost (or payoff) with respect to the player's own action is zero, indicating a local optimum for each player given the actions of others.

Furthermore, if the cost function $f_i(o)$ is continuous in the overall action vector o and exhibits convexity in the individual action o_i while the actions of other players o_{-i} are held constant, the existence of a Nash equilibrium in the non-cooperative game is assured. This assurance stems from the fact that under these conditions, the game's structure allows for a unique solution where no player can benefit from a unilateral deviation, thus ensuring the stability and predictability of the game's outcome. Meanwhile, the convexity of the function lays the foundation for the effectiveness of the gradient descent method, which will be partially utilized later on.

Lemma 1 ([33]). *If Q is a non-singular M-matrix, then $PQ + Q^T P > 0$ for $P = diag(p_1, \ldots, p_N) > 0$.*

Lemma 2 ([34]). *If $c, d > 0$ and $\gamma > 0$, and $x, y \in \mathbb{R}$, the following inequalities hold:*

$$|x|^c |y|^d \leqslant \frac{c}{c+d}\gamma|x|^{c+d} + \frac{d}{c+d}\gamma^{-\frac{c}{d}}|y|^{c+d} \tag{2}$$

3. Modeling of Fixed-Wing Unmanned Aerial Vehicles

Assume that the UAVs are labeled as $i = 1, 2, 3, \ldots, N$, where N represents the number of UAVs in the formation, and all the fixed-wing UAVs are flying at the same altitude. The typical kinematic model for the i-th UAV can be expressed as follows:

$$\dot{x}_i = v_i \cos \theta_i, \ \dot{y}_i = v_i \sin \theta_i, \ \dot{\theta}_i = \omega_i$$
$$\dot{v}_i = u_{i,1}, \ \dot{\omega}_i = u_{i,2} \tag{3}$$

The variables x_i, y_i, and θ_i denote the horizontal coordinate, vertical coordinate, and heading angle of the i-th UAV, respectively. The linear velocity and angular velocity of the UAV are represented by v_i and ω_i, respectively. The control inputs for these dynamics are given by $u_{i,1}$ and $u_{i,2}$.

As described in the paper [30] and illustrated in Figure 1, we define $x_{hi} = x_i + l_i \cos \theta_i$ and $y_{hi} = y_i + l_i \sin \theta_i$, where (x_i, y_i), indicated by the red dot, is the centroid between the head and tail of the i-th UAV, and l_i represents the distance from the centroid to the head of the i-th UAV. As a result, it follows from (3) that

$$\dot{x}_{hi} = \dot{x}_i - l_i \omega_i \sin \theta = v_i \cos \theta_i - l_i \omega_i \sin \theta_i$$
$$\ddot{x}_{hi} = u_{i,1} \cos \theta_i - v_i \omega_i \sin \theta_i - l_i \omega_i^2 \cos \theta_i - l_i u_{i,2} \sin \theta_i$$
$$\dot{y}_{hi} = \dot{y}_i + l_i \omega_i \cos \theta_i = v_i \sin \theta_i + l_i \omega_i \cos \theta_i \tag{4}$$
$$\ddot{y}_{hi} = u_{i,1} \sin \theta_i + v_i \omega_i \cos \theta_i - l_i \omega_i^2 \sin \theta_i + l_i u_{i,2} \cos \theta_i$$

Defining $\ddot{x}_{hi} = u_{x,i}$ and $\ddot{y}_{hi} = u_{y,i}$, one obtains

$$\begin{bmatrix} \ddot{x}_{hi} \\ \ddot{y}_{hi} \end{bmatrix} = \begin{bmatrix} \cos \theta_i & -l_i \sin \theta_i \\ \sin \theta_i & l_i \cos \theta_i \end{bmatrix} \begin{bmatrix} u_{i,1} \\ u_{i,2} \end{bmatrix} + \begin{bmatrix} -v_i \omega_i \sin \theta_i - l_i \omega_i^2 \cos \theta_i \\ v_i \omega_i \cos \theta_i - l_i \omega_i^2 \sin \theta_i \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} u_{i,1} \\ u_{i,2} \end{bmatrix} = \begin{bmatrix} \cos \theta_i & -l_i \sin \theta_i \\ \sin \theta_i & l_i \cos \theta_i \end{bmatrix}^{-1} \begin{bmatrix} u_{x,i} + v_i \omega_i \sin \theta_i + l_i \omega_i^2 \cos \theta_i \\ u_{y,i} - v_i \omega_i \cos \theta_i + l_i \omega_i^2 \sin \theta_i \end{bmatrix} \tag{6}$$
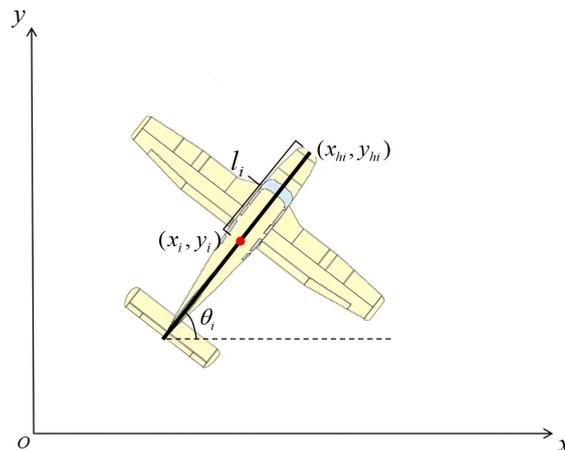

Figure 1. The model of the i-th fixed-wing unmanned aerial vehicle (UAV).

At this point, the following relationships are established:

$$\begin{bmatrix} \dot{x}_{hi} \\ \dot{y}_{hi} \end{bmatrix} = \begin{bmatrix} v_{xi} \\ v_{yi} \end{bmatrix}, \ \begin{bmatrix} \dot{v}_{xi} \\ \dot{v}_{yi} \end{bmatrix} = \begin{bmatrix} u_{x,i} \\ u_{y,i} \end{bmatrix} \tag{7}$$

As a result, $\xi_i = [x_{hi}, y_{hi}]^T$ represents the head coordinates of the UAVs, $v_i = [v_{x,i}, v_{y,i}]^T$ represents the velocities of the UAVs, and $u_i = [u_{x,i}, u_{y,i}]^T$ represents the control signals. Consequently, the following relationships hold:

$$\dot{\xi}_i = v_i, \quad \dot{v}_i = u_i \tag{8}$$

4. Non-Cooperative Game and Nash Equilibrium

4.1. Cost Function

In a non-cooperative game, players aim to minimize their individual cost functions by underscoring the importance of accurately determining these functions.

In the context of UAV formation flying, each UAV pursues two primary objectives. The first objective is to maintain the formation by keeping a specified relative distance from other UAVs, constituting an individual objective. The second objective is to follow the overall reference path ξ_0, which involves maintaining a specified distance relative to ξ_0, fulfilling a collective objective. Here, $\xi_0 = [x_{h0}(t), y_{h0}(t)]^T$ represents a time-dependent function.

To maintain formation, we expect each UAV to maintain a fixed relative position with respect to the others. We defined d_{ij} represents the desired coordinate difference between the head of the i-th UAV and the head of the j-th UAV. To achieve this objective, we aim to minimize

$$\sum_{j=1}^{N} a_{ij} ||\xi_i - \xi_j - d_{ij}||^2 \tag{9}$$

where a_{ij} is the element of the adjacency matrix A and it represents the interaction between the i-th and j-th UAVs.

In order for the UAV formation to track a given mission trajectory as a whole, we expect the UAVs with individual targets to maintain a fixed relative position with respect to the mission trajectory point $\xi_0(t)$. Let p_i be the desired coordinate difference between the head of the i-th UAV and ξ_0. To achieve this goal for agent V_i with independent objectives, we aim to minimize

$$\sum_{j=1}^{N} \alpha_i ||\xi_i - \xi_0 - p_i||^2 \tag{10}$$

where $\alpha_i = 1$ when $V_i \in V_s$ and $\alpha_i = 0$ otherwise. Integrating these two objectives, the cost function $J_i(\xi)$ for the i-th UAV is crafted as a balanced trade-off between formation maintenance and path tracking. The first component ensures that each UAV contributes to the formation's cohesion by minimizing deviations from the desired inter-UAV distances. The second component focuses on the collective objective of tracking the reference path, with select UAVs assuming primary responsibility for this task. Consequently, the cost function $J_i(\xi)$ of the i-th UAV can be designed as follows:

$$J_i(\xi) = \frac{1}{2} \sum_{j=1}^{N} a_{ij} ||\xi_i - \xi_j - d_{ij}||^2 + \frac{\alpha_i}{2} ||\xi_i - \xi_0 - p_i||^2 \tag{11}$$

where $\xi = [\xi_1^T, \ldots, \xi_N^T]^T$, V_s denotes the set of UAVs with independent objectives, and ξ_0 is defined as the given trajectory. By optimizing this cost function, the UAVs within the formation are able to dynamically adjust their positions in response to changes in the environment, ensuring that the formation moves in harmony towards its designated destination with precision and efficiency.

4.2. Nash Equilibrium

Previously, we concluded that in an unconstrained non-cooperative game among players, at the Nash equilibrium, $\frac{\partial f_i(o^*)}{\partial o_i} = 0$ holds. If $f_i(o)$ is continuous in o and convex in o_i with fixed o_{-i}, then a Nash equilibrium exists for the non-cooperative game.

Currently, N UAVs are interconnected as represented by the graph G, with A as the adjacency matrix and L as the Laplacian matrix of the graph. Taking the partial derivative of $J_i(\xi)$ with respect to ξ_i, we derive that

$$\Psi_i(\xi) = \frac{\partial J_i(\xi)}{\partial \xi_i} = \sum_{j=1}^{N} a_{ij}(\xi_i - \xi_j - d_{ij}) + \alpha_i(\xi_i - \xi_0 - p_i). \tag{12}$$

We define $\Psi(\xi) = [\Psi_1(\xi)^T, \ldots, \Psi_N(\xi)^T]^T$, $\Lambda = \text{diag}(\alpha_1, \ldots, \alpha_N)$, and $Q = L + \Lambda$. Let $\alpha = [\alpha_1, \ldots, \alpha_N]^T$. Consequently, we have the following:

$$\Psi(\xi) = (Q \otimes I_3)\xi - d - \alpha \otimes \xi_0, \tag{13}$$

where $\xi = [\xi_1^T, \ldots, \xi_N^T]^T$, $d = [d_1^T, \ldots, d_N^T]^T$, and $d_i = \sum_{j=1}^{N} a_{ij}d_{ij} + \alpha_i p_i$. Given that Q is non-singular, when $\Psi(\xi) = 0_{3N}$, a unique Nash equilibrium state exists, which is denoted as $\xi^* = [\xi_1^{*T}, \ldots, \xi_N^T]^T$. This state can be expressed as follows:

$$\begin{aligned} \xi^* &= (Q \otimes I_3)^{-1}(d + \alpha \otimes \xi_0) \\ &= (Q \otimes I_3)^{-1}d + 1_N \otimes \xi_0 \end{aligned} \tag{14}$$

5. Design of Disturbance Observer Based on Homogeneous Differentiation

The Disturbance Observer (DOB) represents an advanced control strategy specifically tailored to estimate and subsequently compensate for external disturbances that may arise within a system. We will design a DOB based on homogeneous differentiation to observe the system disturbances and compensate for them within the controller. Additionally, we will prove the convergence of this disturbance observer to ensure its stability and effectiveness.

Definition 1. *We define*

$$sig^\alpha(x) = sign(x)|x|^\alpha \tag{15}$$

where $0 \leq \alpha \leq 1$, and sign denote the sign function.

We assume that the disturbances in the system are denoted as $\kappa = [\kappa_1; \kappa_2]$, and under these conditions, for the i-th UAV:

$$\begin{aligned} \dot{x}_i &= v_i \cos\theta_i, \ \dot{y}_i = v_i \sin\theta_i, \ \dot{\theta}_i = \omega_i \\ \dot{v}_i &= u_{i,1} + \kappa_1, \ \dot{\omega}_i = u_{i,2} + \kappa_2 \end{aligned} \tag{16}$$

Assumption 1. *The disturbance κ acting on the system is bounded, meaning that there exists an $\iota > 0$ such that $|\kappa| < \iota$.*

Theorem 1 ([23]). *The exponent power terms of the homogeneous differential observer possess a degree of flexibility, allowing them to be tuned within a predefined range. This adaptability enhances the performance and versatility of the observer. The expression below exemplifies a specific form of the homogeneous differential observer, showcasing its structural characteristics:*

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 - m\,\text{sig}^{1+\alpha}(\hat{x}_1 - g) \\ \dot{\hat{x}}_2 &= -n\,\text{sig}^{1+2\alpha}(\hat{x}_1 - g) \end{aligned} \tag{17}$$

where $-\frac{1}{2} < \alpha < 0$, $m > 0$, $n > 0$, $\hat{x}_1$ and $\hat{x}_2$ are, respectively, the estimated values of g and $\dot{g}$. If

$$|\ddot{g}(t)| < C$$

$$m \geq \left(2^{\frac{1}{1+\alpha}-1}(1 + (2\lambda_1 + 2)\lambda_2 + \lambda_3)n\right)^{\frac{1}{2}} \tag{18}$$

$$n > C$$

where C is a constant, $\lambda_1 = \frac{1}{1+\alpha}\left(\frac{1-2\alpha}{2+\alpha} + \frac{3+4\alpha}{2+\alpha}2^{\frac{1}{1+\alpha}}\right)$, $\lambda_2 = \frac{1+2\alpha}{2+\alpha}2^{\frac{-\alpha}{1+\alpha}}\left(2^{\frac{1}{1+\alpha}}\frac{1-\alpha}{2+\alpha}\right)^{\frac{1-\alpha}{1+2\alpha}}$,

$\lambda_3 = \frac{1}{1+\alpha}\frac{3+4\alpha}{2+\alpha}2^{\frac{1}{1+\alpha}}$.

Then, the aforementioned homogeneous differential observer exhibits a property of finite-time convergence. This signifies that the estimation error generated by the observer rapidly diminishes and converges precisely to zero within a predetermined, finite time interval. In contrast, it does not follow the traditional asymptotic convergence pattern where the error tends to zero as time approaches infinity. The finite-time convergence feature ensures that the observer's estimates accurately reflect the actual values in a timely and efficient manner.

Lemma 3. *By using Theorem 1, we can draw the following conclusion: A homogeneous differential disturbance observer in the following form*

$$\dot{\hat{v}} = -msig^{(}1+\alpha)(\hat{v} - v) + \hat{\kappa} - \kappa + u$$
$$\dot{\hat{\kappa}} = -nsig^{(}1+2\alpha)(\hat{v} - v), \tag{19}$$

can track the internal disturbances within the system within a finite time. That is, $\hat{\kappa}$ will converge to κ within a finite time. Where $-\frac{1}{2} < \alpha < 0, m > 0, n > 0$, $\hat{\kappa}$ *is the output of the disturbance observer.*

Proof. Define the Estimated value x_1 and x_2 of this homogeneous differential disturbance observer as

$$x_1 = \hat{v}$$
$$x_2 = \hat{\kappa} \tag{20}$$

Therefore, we can obtain the dynamic equation of the estimation error as follows:

$$\dot{x_1} = -msig^{1+\alpha}(x_1 - v) + x_2$$
$$\dot{x_2} = -nsig^{1+2\alpha}(x_1 - v) \tag{21}$$

If the homogeneous differential observer is transformed into this particular form, it becomes readily apparent that it fulfills the necessary criteria stipulated by Theorem 1. Consequently, the estimation $\hat{\kappa}$ of the disturbance κ generated by this tailored observer exhibits the property of finite-time convergence. This implies that the estimation error converges precisely to zero within a finite time interval, ensuring the timely and accurate reflection of the actual disturbance value. $\square$

After offsetting the estimated value of the disturbance against the actual disturbance for the *i*-th UAV is as follows:

$$\dot{x}_i = v_i \cos\theta_i, \ \dot{y}_i = v_i \sin\theta_i, \ \dot{\theta}_i = \omega_i$$
$$\dot{v}_i = (u_{i,1} - \hat{\kappa_1}) + \kappa_1, \dot{\omega}_i = (u_{i,2} - \hat{\kappa_2}) + \kappa_2 \tag{22}$$

6. Implementation of Fixed-Wing UAV Formation Control Based on Non-Cooperative Game

We have presented the designed controller and proven its feasibility.

Theorem 2. *Assuming that the objective of the UAVs is to minimize the cost function $J_i(\xi)$, if a gradient control strategy is designed as follows:*

$$u_i = \ddot{\xi}_0 - \dot{\Psi}_i(\xi) - c(\Psi_i(\xi) + v_i - \dot{\xi}_0) \tag{23}$$

then the UAVs will gradually converge to the Nash equilibrium state, where c is a coefficient to be determined.

Proof. Let $v = [v_1^T, \ldots, v_N^T]^T$ and $u = [u_1^T, \ldots, u_N^T]^T$. Then, the following relationships hold:

$$\dot{\xi} = v, \quad \dot{v} = u \tag{24}$$

Feasibility is proved using backstepping theory [35]. Initially, let $\varepsilon_1 = \xi - \xi^*$, and choose the Lyapunov function as follows:

$$V_1 = \varepsilon_1^T (P \otimes I_3) \varepsilon_1 \tag{25}$$

where P is a positive definite matrix. It follows that

$$
\begin{aligned}
\dot{V}_1 =& (\xi - \xi^*)^T (P \otimes I_3)(v - 1_N \otimes \dot{\xi}_0) \\
&+ (v - 1_N \otimes \dot{\xi}_0)^T (P \otimes I_3)(\xi - \xi^*)
\end{aligned} \tag{26}
$$

Using the fact that $\Psi(\xi^*) = 0_{3N}$, we have

$$
\begin{aligned}
\dot{V}_1 =& (\xi - \xi^*)^T (P \otimes I_3)(\Psi(\xi) + v - 1_N \otimes \dot{\xi}_0) \\
&- (\xi - \xi^*)^T (P \otimes I_3)(\Psi(\xi) - \Psi(\xi^*)) \\
&+ (\Psi(\xi) + v - 1_N \otimes \dot{\xi}_0)^T (P \otimes I_3)(\xi - \xi^*) \\
&- (\Psi(\xi) - \Psi(\xi^*))^T (P \otimes I_3)(\xi - \xi^*)
\end{aligned} \tag{27}
$$

From $\Psi(\xi) - \Psi(\xi) = (Q \otimes I_3)(\xi - \xi)$, it can be derived that

$$
\begin{aligned}
\dot{V}_1 =& (\xi - \xi^*)^T (P \otimes I_3)(\Psi(\xi) + v - 1_N \otimes \dot{\xi}_0) \\
&- (\xi - \xi^*)^T (PQ \otimes I_3)(\xi - \xi^*) \\
&+ (\Psi(\xi) + v - 1_N \otimes \dot{\xi}_0)^T (P \otimes I_3)(\xi - \xi^*) \\
&- (\xi - \xi^*)^T (Q^T P \otimes I_3)(\xi - \xi^*)
\end{aligned} \tag{28}
$$

Given that $-PQ - Q^T P < 0$, we define $\lambda = \lambda_{\min}(PQ + Q^T P)$, $p_{max} = \max(p_1, \ldots, p_N)$, and $p_{min} = \min(p_1, \ldots, p_N)$, which leads to the following inequality:

$$
\begin{aligned}
\dot{V}_1 \leq& -\lambda ||\xi - \xi^*||^2 \\
&+ 2p_{max} ||\xi - \xi^*|| \cdot ||\Psi(\xi) + v - 1_N \otimes \dot{\xi}_0||
\end{aligned} \tag{29}
$$

Let $v^* = 1_N \otimes \dot{\xi}_0 - \Psi(\xi)$. We can obtain the following:

$$
\begin{aligned}
\dot{V}_1 \leq& -\lambda ||\xi - \xi^*||^2 \\
&+ 2p_{\max} ||\xi - \xi^*|| \cdot ||\Gamma(\xi) + v + v^* - 1_N \otimes \dot{\xi}_0|| \\
\leq& -\lambda ||\xi - \xi^*||^2 + 2p_{\max} ||\xi - \xi^*|| \cdot ||v - v^*|| \\
\leq& -\lambda ||\xi - \xi^*||^2 + ||\xi - \xi^*||^2 \\
&+ p_{\max}^2 ||\xi - \xi^*|| \cdot ||v - v^*||
\end{aligned} \tag{30}
$$

Defining $\varepsilon_2 = v - v_d = v - v^*$, it can be deduced that

$$\dot{\varepsilon}_2 = (Q \otimes I_3)v - \alpha \otimes \dot{\xi}_0 + \dot{v} - 1_N \otimes \ddot{\xi}_0 \tag{31}$$

We propose a Lyapunov function for the second part of the system:

$$V_2 = \frac{1}{2} \varepsilon_2^T \varepsilon_2 \tag{32}$$

The derivative of V_2 is

$$\dot{V}_2 = (\nu - \nu^*)^T (u - \dot{\nu}^*) \tag{33}$$

If

$$\begin{aligned}
u &= \dot{\nu}^* - c(\nu - \nu^*) \\
&= 1_N \otimes \ddot{\xi}_0 - \dot{\Psi}(\xi) - k(\Psi(\xi) + \nu - 1_N \otimes \dot{\xi}_0)
\end{aligned} \tag{34}$$

then

$$\dot{V}_2 = -c(\nu - \nu^*)^T (\nu - \nu^*) \leq 0 \tag{35}$$

where c is a positive real number. Choose the total Lyapunov function

$$V_1 + V_2 = \varepsilon_1^T (P \otimes I_3) \varepsilon_1 + \frac{1}{2} \varepsilon_2^T \varepsilon_2 \geq 0 \tag{36}$$

Since $P \otimes I_3$ is a positive definite matrix, then $\lambda_{\max}(P \otimes I_3) > 0$, which implies that

$$\varepsilon_1^T (P \otimes I_3) \varepsilon_1 \leq \lambda_{\max}(P \otimes I_3) \varepsilon_1^T \varepsilon_1 \tag{37}$$

As a result

$$\begin{aligned}
V_1 + V_2 &\leq \lambda_{\max}(P \otimes I_3) \varepsilon_1^T \varepsilon_1 + \frac{1}{2} \varepsilon_2^T \varepsilon_2 \\
&\leq \lambda_1 (\varepsilon_1^T \varepsilon_1 + \varepsilon_2^T \varepsilon_2)
\end{aligned} \tag{38}$$

where $\lambda_1 = \max\{\lambda_{\max}(P \otimes I_3), \frac{1}{2}\}$. At the same time, we have

$$\begin{aligned}
\dot{V}_1 + \dot{V}_2 &= (\xi - \xi^*)^T (P \otimes I_3)(\nu - 1_N \otimes \dot{\xi}_0) \\
&\quad + (\nu - 1_N \otimes \dot{\xi}_0)^T (P \otimes I_3)(\xi - \xi^*) \\
&\quad + (\nu - \nu^*)^T (u - \dot{\nu}^*) \\
&\leq -\lambda \|\xi - \xi^*\|^2 + \|\xi - \xi^*\|^2 \\
&\quad + p_{\max}^2 \|\xi - \xi^*\| \cdot \|\nu - \nu^*\| \\
&\quad - c(\nu - \nu^*)^T (\nu - \nu^*)
\end{aligned} \tag{39}$$

By Lemma 2, $p_{max}^2 \|\xi - \xi^*\| \cdot \|\nu - \nu^*\|$ can be calculated by the following:

$$\begin{aligned}
p_{max}^2 \|\xi - \xi^*\| \cdot \|\nu - \nu^*\| & \\
&\leq \frac{p_{max}^2 \gamma}{2} \|\xi - \xi^*\|^2 + \frac{\gamma^{-1} p_{max}^2}{2} \|\nu - \nu^*\|^2
\end{aligned} \tag{40}$$

where γ is a positive number. By substituting (40) into (39), we obtain

$$\begin{aligned}
\dot{V}_1 + \dot{V}_2 &\leq -(\lambda - 1 - \frac{p_{\max}^2 \gamma}{2}) \|\xi - \xi^*\|^2 \\
&\quad - (c - \frac{\gamma^{-1} p_{\max}^2}{2}) \|\nu - \nu^*\|^2
\end{aligned} \tag{41}$$

Letting $\gamma = \frac{\lambda}{p_{max}^2} > 0$, we obtain that

$$\dot{V}_1 + \dot{V}_2 \leq -(\frac{\lambda}{2} - 1)(\varepsilon_1^T \varepsilon_1) - (c - \frac{p_{max}^4}{2\lambda})(\varepsilon_2^T \varepsilon_2) \tag{42}$$

Since the parameters c and λ can be adjusted by modifying the controller gains and the matrix P, if selecting appropriate values to satisfy $c \geq \frac{p_{max}^4}{2\lambda}$ and $\lambda \geq 2$, it can be concluded that

$$\dot{V}_1 + \dot{V}_2 \leq -\lambda_2(\varepsilon_1^T \varepsilon_1 + \varepsilon_2^T \varepsilon_2) \tag{43}$$

where $\lambda_2 = \min\left\{\frac{\lambda}{2} - 1, c - \frac{p_{max}^4}{2\lambda}\right\} > 0$. This, combined with (38), leads to the following:

$$\dot{V}_1 + \dot{V}_2 \leq -\lambda_2(\varepsilon_1^T \varepsilon_1 + \varepsilon_2^T \varepsilon_2) \leq -\frac{\lambda_2}{\lambda_1}(V_1 + V_2) \tag{44}$$

Defining $\lambda_3 = \frac{\lambda_2}{\lambda_1} > 0$, we arrive at the conclusion:

$$\dot{V}_1 + \dot{V}_2 \leq -\lambda_3(V_1 + V_2) \tag{45}$$

Since $V_1 + V_2$ is positive definite and $\dot{V}_1 + \dot{V}_2$ is negative definite, we determine that the system is exponentially stable. It follows that $\varepsilon_1, \varepsilon_2 \to 0$ as $t \to \infty$, $\xi \to \xi^*$ while $t \to \infty$. Through the above derivation, it can be concluded that under the given control law, the UAV cluster will gradually converge to the Nash equilibrium state.   □

7. Simulation and Validation

We utilize MATLAB R2023b as our simulation platform to validate the effectiveness of the proposed control algorithm.

In our simulation, we set $N = 4$, representing a cluster of four UAVs. These UAVs are designed to operate in a coordinated manner, following a predetermined overall path. The desired path ξ_0 is configured as a circular trajectory. This choice of trajectory allows us to evaluate the algorithm's performance in a continuous and repetitive pattern.

To visualize the formation and coordination of the four UAVs, we present a schematic diagram (Figure 2), which depicts their relative positions and orientations within the formation. 1, 2, 3, and 4 are their respective identification numbers, corresponding to the first, second, third, and fourth UAVs. The UAVs are arranged in a specific pattern, which ensures optimal coverage and coordination during flight.
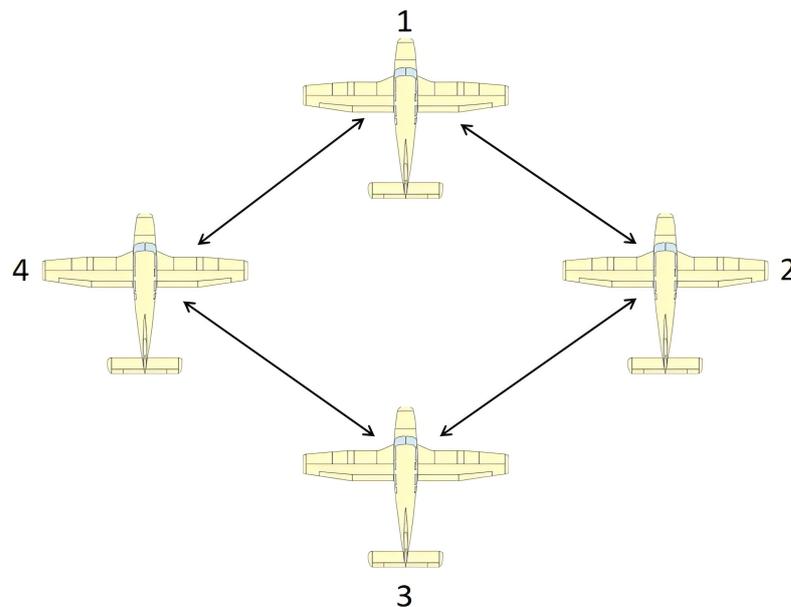
Figure 2. Schematic diagram of the UAV formation with four UAVs.

When the network connectivity relationships between the UAVs are established, the values of the adjacency matrix for the UAV connection graph can be determined. This matrix serves as a fundamental mathematical representation of the connectivity and information flow between the UAVs.

The adjacency matrix for the UAV connection graph is defined as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \tag{46}$$

To accurately quantify the pose differences among the UAVs within a network, we introduce a matrix representation called the pose difference matrix, which denoted by D^*, provides a structured view of the variations in position and orientation among the UAVs. The matrix D^* is structured as follows:

$$D^* = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \\ d_{31} & d_{32} & d_{33} & d_{34} \\ d_{41} & d_{42} & d_{43} & d_{44} \end{bmatrix} = \begin{bmatrix} 0 & 10 & 0 & 10 \\ 0 & 0 & 10 & 10 \\ -10 & 0 & -10 & 0 \\ 0 & 0 & 10 & 10 \\ 0 & 10 & 0 & 10 \\ -10 & -10 & 0 & 0 \\ -10 & 0 & -10 & 0 \\ -10 & -10 & 0 & 0 \end{bmatrix} \tag{47}$$

To accurately determine the pose differences between the UAVs and the reference path points, we first establish a matrix representation of these differences. Specifically, we define the matrix $P^* = [p_1^T, p_2^T, p_3^T, p_4^T]^T$, where p_i^T represents the transpose of the pose difference vector for the i-th UAV with respect to the corresponding reference path point. In this example, $P^* = [-5, -5, 5, -5, -5, 5, 5, 5]^T$.

We utilize the mathematical formulation represented by Equation (14) to calculate the Nash equilibrium trajectory for a formation of four UAVs, which symbolizes a state where the actions of each UAV are balanced with those of the others, resulting in a stable formation configuration.

By applying Equation (14), we determine the optimal trajectories for each UAV in the formation. The results of this calculation are illustrated in Figure 3, where the Nash equilibrium trajectories of the four UAVs are displayed, showing how they coordinate their movements to follow the desired path while maintaining the required formation.

In Figure 3, the prominent solid black line serves as a representation of the intended path that the UAV formation is required to follow. Consequently, the calculated Nash equilibrium trajectories for each individual UAV within the formation are precisely depicted. Specifically, the Nash equilibrium trajectory for UAV 1 is illustrated in red, UAV 2 in pink, UAV 3 in green, and UAV 4 in blue.

Additionally, we introduce the disturbance $\kappa = [\kappa_1; \kappa_2] = [4 \times \sin(t+1); 3 \times \cos(2t)]$, which represents a significant disturbance with relatively large amplitude and varying speed. It can be observed that during the period when the disturbance is acting, the estimated value $\hat{\kappa}$ from the homogeneous differential disturbance observer rapidly converges to the actual value κ. By incorporating this estimated disturbance value $\hat{\kappa}$ into the control strategy by subtracting it from the control input $u_i = [u_{i,1}; u_{i,2}]$, we effectively counteract the disturbance's detrimental effects on the system. As evident from the subsequent graphical representation Figure 4, the homogeneous differential disturbance observer plays a pivotal role in minimizing the disturbance's impact, rendering its influence on the system virtually negligible.

Upon activation of the control signal (23), it begins to influence the four UAVs. After a brief period of adjustment and alignment, the UAVs successfully achieve a cohesive

formation. This unified entity then proceeds to follow the predefined desired trajectory. As depicted in Figure 5 and Figure 6, the UAVs maintain their relative positions within the formation, ensuring synchronized movement along the intended path.
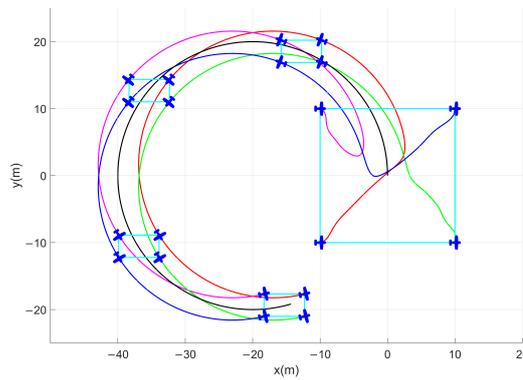
The control system diagram is shown in Figure 7.
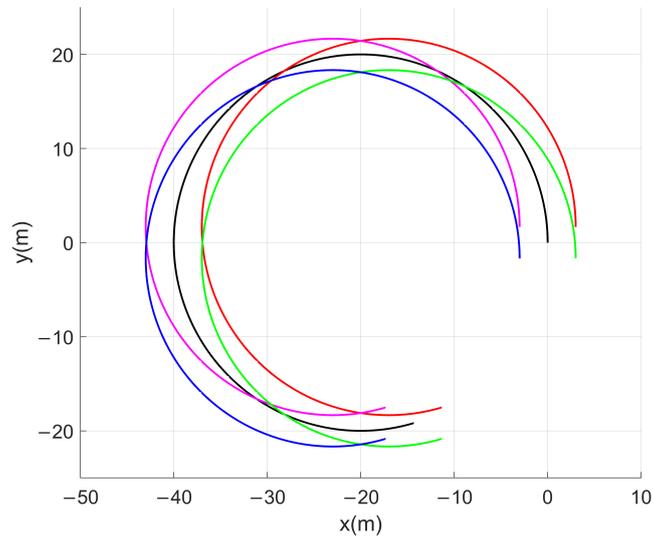
Figure 3. The computed trajectories of four UAVs at the Nash equilibrium. The trajectories of the UAVs are colored red for the first UAV, pink for the second UAV, green for the third UAV, and blue for the fourth UAV.
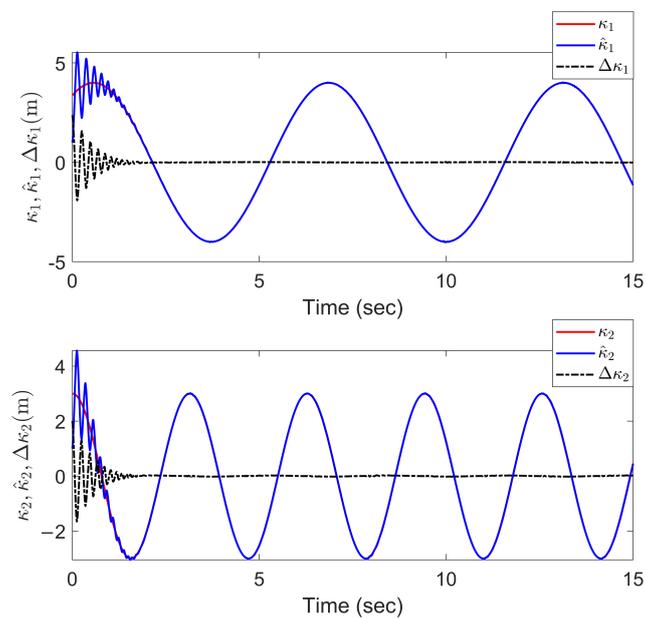
Figure 4. Actual value of applied disturbance, estimated value of the disturbance by the disturbance observer and estimation error.

Figure 5. Trajectories of four UAVs in a simulation environment. The trajectories of the UAVs are colored red for the first UAV, pink for the second UAV, green for the third UAV, and blue for the fourth UAV.
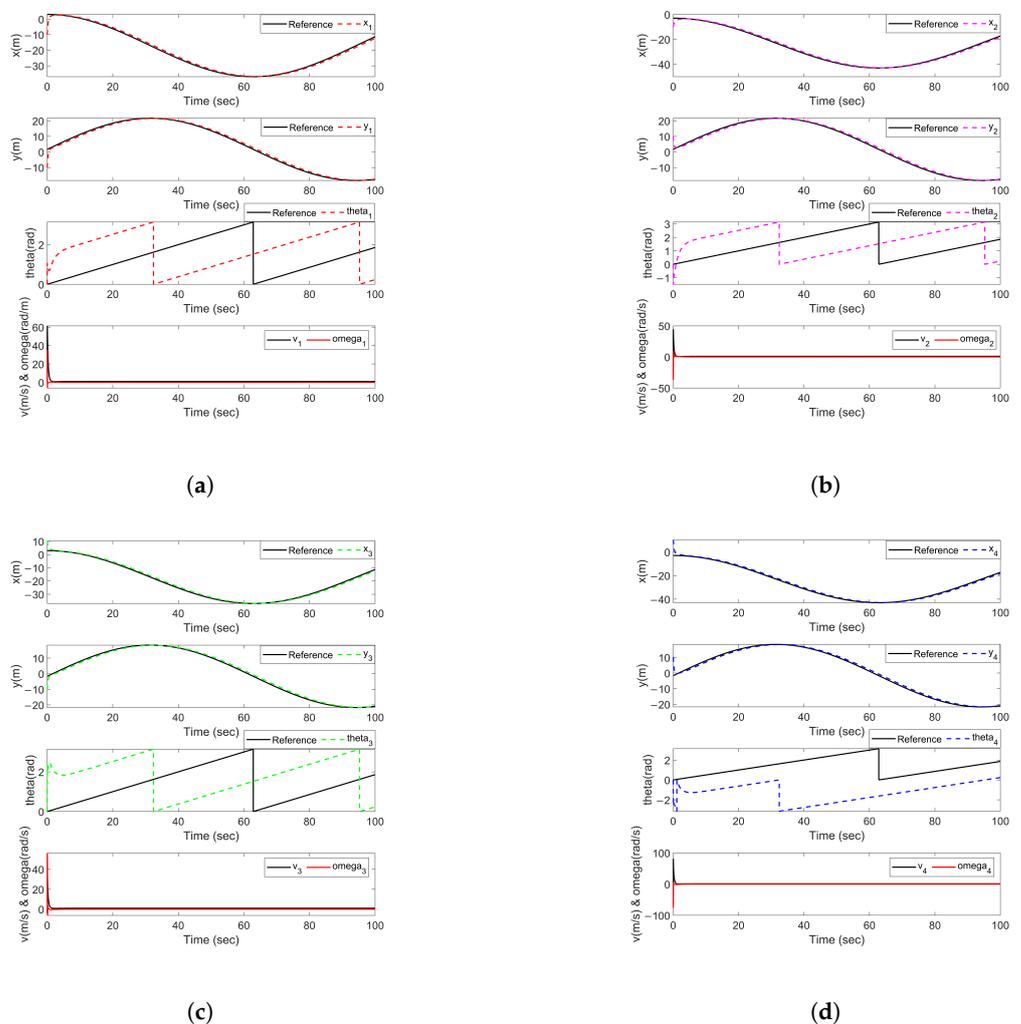
Figure 6. The pose variation of UAVs during the simulation process. (**a**) The pose variation of UAV 1 during the simulation process. (**b**) The pose variation of UAV 2 during the simulation process. (**c**) The pose variation of UAV 3 during the simulation process. (**d**) The pose variation of UAV 4 during the simulation process.
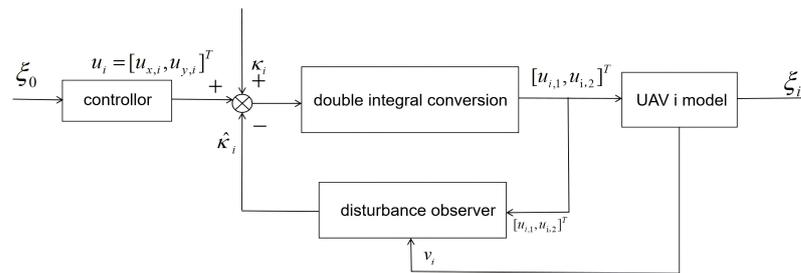
Figure 7. The control system diagram of four UAVs.

After careful analysis of the collected data, we conclude that a fixed-wing UAV formation will reach the desired Nash equilibrium of the specified non-cooperative game within a set time, assuming its initial state is reasonable. Once the Nash equilibrium is achieved, the formation will maintain its structure and adhere to the planned trajectory, ensuring smooth and coordinated flight.

8. Conclusions

Given the widespread application of fixed-wing UAV formations and the communication and control challenges encountered in the process, this paper proposes a distributed fixed-wing UAV formation control method based on non-cooperative game theory. Firstly, we assumed a planar task environment and established a UAV model based on this assumption. We further transformed this model into a double integrator form, which reduces the complexity of controller design. Next, we implemented a homogeneous differential disturbance observer based on the UAV model to enhance the UAV's robustness against disturbances generated by the system itself during operation. Considering the two objectives in UAV formation tasks: maintaining formation and following the mission trajectory, we designed a cost function that balances both aspects. The UAV's goal is to minimize its own cost function. Therefore, utilizing non-cooperative game theory, we identified the Nash equilibrium solution. Subsequently, we employed the backstepping method to design an appropriate controller, which enables the states of the multi-UAV system to ultimately converge to the Nash equilibrium point. Finally, simulations were conducted to demonstrate the feasibility of the proposed algorithm.

References

1. Kopfstedt, T.; Mukai, M.; Fujita, M.; Ament, C. Control of formations of UAVs for surveillance and reconnaissance missions. *IFAC Proc. Vol.* **2008**, *41*, 5161–5166. [CrossRef]
2. Duan, H.; Qiu, H.; Chen, L.; Wei, C. Prospects on unmanned aerial vehicle autonomous swarm technology. *Sci. Technol. Rev.* **2018**, *36*, 90–97.
3. Wu, J.; Wang, H.; Li, S.; Liu, S. Distributed Adaptive Path-Following Control for Distance-Based Formation of Fixed-Wing UAVs under Input Saturation. *Aerospace* **2023**, *10*, 768. [CrossRef]

4. Wu, G.; Xu, T.; Sun, Y.; Zhang, J. Review of multiple unmanned surface vessels collaborative search and hunting based on swarm intelligence. *Int. J. Adv. Robot. Syst.* **2022**, *19*. [CrossRef]
5. Wang, G.; Luo, H.; Hu, X.; Ma, H.; Yang, S. Fault-tolerant communication topology management based on minimum cost arborescence for leader–follower UAV formation under communication faults. *Int. J. Adv. Robot. Syst.* **2017**, *14*. [CrossRef]
6. Xiao, Y.; Zhou, L. Dynamic Modelling and Numerical Simulation of Formation Control for Intelligent Multi-Agent System with Target Geometric Configuration. *Appl. Math. Comput.* **2023**, *444*, 127826. [CrossRef]
7. Chen, Q.; Wang, T.; Jin, Y.; Wang, Y.; Qian, B. A UAV Formation Control Method Based on Sliding-Mode Control under Communication Constraints. *Drones* **2023**, *7*, 231. [CrossRef]
8. Du, Z.; Qu, X.; Shi, J.; Lu, J. Formation control of fixed-wing UAVs with communication delay. *ISA Trans.* **2024**, *146*, 154–164. [CrossRef]
9. Suo, W.; Wang, M.; Zhang, D.; Qu, Z.; Yu, L. Formation control technology of fixed-wing UAV swarm based on distributed ad hoc network. *Appl. Sci.* **2022**, *12*, 535. [CrossRef]
10. Guinaldo, M.; Sánchez, J.; Dormido, S. Distributed adaptive control of linear multi-agent systems with event-triggered communications. *Appl. Math. Comput.* **2016**, *274*, 195–207. [CrossRef]
11. Cao, Y.; Ren, W.; Egerstedt, M. Distributed containment control with multiple stationary or dynamic leaders in fixed and switching directed networks. *Automatica* **2012**, *48*, 1586–1597. [CrossRef]
12. Wei, H.; Sun, Q.; Chen, J.; Shi, Y. Robust distributed model predictive platooning control for heterogeneous autonomous surface vehicles. *Control Eng. Pract.* **2021**, *107*, 104655. [CrossRef]
13. Huang, J.; Chen, C.; Fang, H.; Chen, J.; Dou, L. On consensus of multiple high-order uncertain systems based on distributed backstepping framework. *Int. J. Adv. Robot. Syst.* **2014**, *11*, 120. [CrossRef]
14. Safaei, A.; Mahyuddin, M.N. Distributed adaptive model-free cooperative control for a network of generic unknown nonlinear systems. *Int. J. Adv. Robot. Syst.* **2018**, *15*. . [CrossRef]
15. Yan, S.; Feng, J.; Pan, F. A Distributed Task Allocation Method for Multi-UAV Systems in Communication-Constrained Environments. *Drones* **2024**, *8*, 342. [CrossRef]
16. Zhang, Y.; Wu, D.; Cheng, P.; Wu, W.; Zhang, W. Robust adaptive fault-tolerant control for path maneuvering of autonomous surface vehicles with actuator faults based on the noncooperative game strategy. *Ocean Eng.* **2024**, *292*, 116541. [CrossRef]
17. Liang, S.; Yi, P.; Hong, Y. Distributed Nash equilibrium seeking for aggregative games with coupled constraints. *Automatica* **2017**, *85*, 179–185. [CrossRef]
18. Ye, M.; Hu, G.; Lewis, F.L.; Xie, L. A Unified Strategy for Solution Seeking in Graphical *N*-Coalition Noncooperative Games. *IEEE Trans. Autom. Control* **2019**, *64*, 4645–4652. [CrossRef]
19. Lu, K.; Jing, G.; Wang, L. Distributed Algorithms for Searching Generalized Nash Equilibrium of Noncooperative Games. *IEEE Trans. Cybern.* **2019**, *49*, 2362–2371. [CrossRef]
20. Ohnishi, K. A New Servo Method in Mechatronics. *Trans. Jpn. Soc. Electr. Eng. D* **1987**, *177*, 83–86.
21. Liu, X.; Wang, M.; Zheng, Y.; Wang, X. Fuzzy PI vibration suppression control strategy for space double flexible telescopic manipulator with fractional disturbance observer. *Aerosp. Sci. Technol.* **2024**, *155*, 109579. [CrossRef]
22. Hu, X.; Chen, J.; Lv, C.; Zhang, Z.; Wang, Y. Finite-time trajectory tracking control for unmanned ground vehicle based on finite-time disturbance observer. *J. Frankl. Inst.* **2024**, *361*, 107222. [CrossRef]
23. Chen, W.; Du, H. A second-order discrete-time super-twisting control algorithm based on homogeneous system theory. *Control Theory Technol.* **2022**, *39*, 761–769.
24. Frazzoli, E.; Dahleh, M.; Feron, E. Trajectory tracking control design for autonomous helicopters using a backstepping algorithm. In Proceedings of the American Control Conference, Chicago, IL, USA, 28–30 June 2000.
25. Cabecinhas, D.; Cunha, R.; Silvestre, C. A globally stabilizing path following controller for rotorcraft with wind disturbance rejection. *IEEE Trans. Control Syst. Technol.* **2014**, *23*, 708–714. [CrossRef]
26. Zhang, J.; Yan, J.; Zhang, P. Multi-UAV formation control based on a novel back-stepping approach. *IEEE Trans. Veh. Technol.* **2020**, *69*, 2437–2448. [CrossRef]
27. Chang, Y.; Feng, Z.; Zhang, X. Backstepping-based adaptive fuzzy tracking control for pure-feedback nonlinear multi-agent systems. *Int. J. Syst. Sci.* **2024**, *55*, 1584–1595. [CrossRef]
28. Zheng, X.; Yang, J.; Jiang, J.; Wang, X. Gradient Descent Algorithm Based Backstepping NN Control of a Perturbed Quad-Rotor UAV: Theory and Experiments. In Proceedings of the 42nd Chinese Control Conference, Tianjin, China, 24–26 July 2023.
29. Wang, F.; Gao, Y.; Zhou, C.; Zong, Q. Disturbance observer-based backstepping formation control of multiple quadrotors with asymmetric output error constraints. *Appl. Math. Comput.* **2022**, *415*, 126693. [CrossRef]
30. Cheng, Y.; Jia, R.; Du, H.; Wen, G.; Zhu, W. Robust finite-time consensus formation control for multiple nonholonomic wheeled mobile robots via output feedback. *Int. J. Robust Nonlinear Control* **2018**, *28*, 2082–2096. [CrossRef]
31. Ren, W.; Beard, R.W.; Atkins, E.M. Information consensus in multivehicle cooperative control. *IEEE Control Syst. Mag.* **2007**, *27*, 71–82.
32. Furuyama, S.; Nakai, T. The construction of subjective games by motive distributions in n -person non-cooperative game. *J. Inf. Optim. Sci.* **2004**, *25*, 533–541. [CrossRef]
33. Peng, Z.; Wang, J.; Wang, D.; Han, Q.L. An Overview of Recent Advances in Coordinated Control of Multiple Autonomous Surface Vehicles. *IEEE Trans. Ind. Inform.* **2021**, *17*, 732–745. [CrossRef]

34. Qian, C.; Lin, W. A continuous feedback approach to global strong stabilization of nonlinear systems. *IEEE Trans. Autom. Control* **2001**, *46*, 1061–1079. [CrossRef]
35. Zhang, X.; Jiang, K. Backstepping-based adaptive control of underactuated AUV subject to unknown dynamics and zero tracking errors. *Ocean Eng.* **2024**, *302*, 117640. [CrossRef]