SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

May 2026, Vol. 69, Iss. 5, 152205:1–152205:20
https://doi.org/10.1007/s11432-025-4765-9

Distributed fault-tolerant control for multi-agent pursuit-evasion games under communication link faults

Wenjing HOU, Li LIANG & Youqing WANG*

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

Received 31 July 2025/Revised 28 September 2025/Accepted 27 October 2025/Published online 23 March 2026

Abstract This study investigates fault-tolerant control (FTC) strategies for multi-pursuer single-evader pursuit-evasion (PE) games subject to communication link faults. To address the challenges posed by time-varying and unknown topological weights caused by these faults, a distributed adaptive observer with dynamic gain adjustment is designed to enable each pursuer to estimate the evader's state accurately, even under unreliable or corrupted links. Based on the estimated states, performance index functions are constructed by incorporating relative state errors and control efforts, thereby formulating the game dynamics. To approximate the solution of the underlying Hamilton-Jacobi-Isaacs (HJI) equation, which characterizes the Nash equilibrium of the differential game, a single-critic neural network (NN) is employed for real-time pursuit strategy updates. This approach significantly reduces computational complexity while ensuring strategic optimality. A Lyapunov-based analysis rigorously establishes the convergence of the observer estimation errors and guarantees that the relative state between each pursuer and the evader converges to zero, ensuring successful capture despite communication faults. Finally, simulation results based on a nonlinear unicycle dynamics model validate the effectiveness and robustness of the proposed control framework in the presence of communication link faults.

Keywords pursuit-evasion game, fault-tolerant control, communication link faults, distributed adaptive observer, critic neural network

Citation Hou W J, Liang L, Wang Y Q. Distributed fault-tolerant control for multi-agent pursuit-evasion games under communication link faults. Sci China Inf Sci, 2026, 69(5): 152205, https://doi.org/10.1007/s11432-025-4765-9

1 Introduction

Multi-agent pursuit-evasion (MPE) games represent a fundamental and dynamic branch of game theory, addressing complex scenarios in which multiple pursuers aim to intercept or constrain one or more evaders. These games encapsulate the intricate interplay between adversarial agents and necessitate the design of intelligent strategies that account for practical constraints such as limited resources, environmental uncertainty, and partial observability. As the number of agents increases, the coordination burden among pursuers grows, and evaders must adopt increasingly adaptive escape behaviors. MPE games not only enrich theoretical research in distributed decision-making and differential games but also find wide applications in real-world systems, such as autonomous aerial vehicles [1], unmanned surface vehicles [2], satellite formations [3], and spacecraft control systems [4].

The early research on PE games primarily focused on the basic scenario of a single pursuer and a single evader, as first introduced in [5] and later extended in [6–9]. In this classical setting, typically formulated as a zero-sum differential game, the pursuer aims to minimize a predefined cost function while the evader seeks to maximize it, with each agent making decisions independently. Solutions are often obtained by solving the associated Hamilton-Jacobi-Isaacs (HJI) equations. Subsequent studies expanded to multi-agent configurations, including two-pursuer scenarios [10, 11] and more general cases involving multiple pursuers [12–16]. Among these, the multi-pursuer single-evader setting has garnered significant attention due to its practical applicability and increased theoretical complexity. It demands sophisticated coordination strategies to capture a highly maneuverable evader in dynamic and uncertain environments. However, most existing studies assume ideal communication among agents, neglecting real-world challenges such as communication link faults, which can severely degrade coordination performance by disrupting timely and accurate information exchange.

Effective coordination and strategy synchronization in MPE games heavily rely on continuous and reliable information exchange through communication links. However, in real-world scenarios, such links are often exposed to stochastic failures, electromagnetic interference in cluttered environments, and adversarial cyber-attacks, leading to time-varying and unpredictable variations in topological weights and ultimately communication faults. To address these challenges, various fault-tolerant control (FTC) methods have been proposed. Adaptive synchronization

* Corresponding author (email: wang.youqing@ieee.org)

techniques were developed in [17] to ensure consistent coordination despite faulty communications. Reinforcement learning and adaptive control have been applied in [18, 19] for formation control of heterogeneous vehicles with enhanced robustness to inter-agent communication disruptions. In [20], a reduced-order observer and fuzzy adaptive control were combined into dynamic event-triggered consensus schemes, enabling real-time adaptation. Other studies such as [21] investigated adaptive control strategies for nonlinear agents with dead-zone inputs and event-triggered mechanisms to maintain system performance under faults. Additionally, Refs. [22, 23] discussed adaptive leader-following consensus and synchronization to improve resilience in complex networks. Recent advances also include fixed-time adaptive neural tracking control for heterogeneous nonlinear systems [24]. These methods reflect the increasing demand for robustness and adaptability in distributed control. Despite these developments, most studies do not explicitly consider communication faults in the context of nonlinear MPE games with game-theoretic strategies, which highlights the necessity to develop distributed FTC frameworks tailored to such scenarios.

In practical applications, the dynamics of MPE games are typically governed by nonlinear ordinary differential equations [25,26], which lead to nonlinear HJI equations that are analytically intractable. To address this challenge, approximate or adaptive dynamic programming (ADP) techniques have been developed for solving nonlinear differential game problems [27–32], often relying on neural networks to approximate the value functions. Among these, actor-critic architectures are widely adopted, where separate actor and critic networks are used to learn control policies and value functions, respectively [26, 33]. Although such decentralized structures support distributed strategy learning, they also introduce significant computational overhead. To reduce this burden, single-network adaptive critic designs have been proposed for constrained optimal control problems [34, 35], eliminating the need for an explicit policy network. Despite substantial progress in strategy design for MPE games [36–42], little attention has been paid to the impact of communication link faults in such settings. These faults can severely disrupt Laplacian-based cooperative estimation and compromise the stability of Nash equilibrium strategies, posing fundamental challenges to evader capture in multi-agent games.

Building on the previous discussion, this paper tackles the challenges introduced by communication link faults in MPE games, aiming to bridge the gap between FTC and game-theoretic multi-agent coordination. To this end, a novel control framework is proposed, which integrates a distributed adaptive observer to handle time-varying and uncertain communication topologies. This design enhances both estimation accuracy and system stability under faulty network conditions. The main contributions of this work are summarized as follows.

(1) This work explicitly incorporates communication link faults into both the modeling and control design of PE games, addressing a critical yet underexplored aspect often neglected in existing studies. The proposed framework captures the impact of time-varying topological disruptions, offering a more robust and practical solution for multi-agent coordination.

(2) A distributed adaptive observer with dynamic gain regulation is developed to estimate the evader's state under communication faults. This design ensures estimation accuracy despite intermittent or degraded links, thereby enhancing fault-tolerant performance.

(3) A single-critic neural network (NN) is employed to approximate the Nash equilibrium solution of the differential game. This structure improves computational efficiency while maintaining accuracy, facilitating real-time implementation in nonlinear multi-agent systems.

The remainder of this paper is organized as follows. Section 2 introduces the preliminaries and problem formulation. Section 3 details the main theoretical results. Section 4 provides a simulation example based on a nonlinear unicycle kinematic model to demonstrate the effectiveness of the proposed methods. Finally, Section 5 concludes the paper and discusses potential directions for future research.

2 Preliminaries and problem formulation

2.1 Graph theory

Consider a multi-pursuer single-evader PE game consisting of M pursuers and one evader. The communication topology among pursuers is modeled by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the node set $\mathcal{V} = \{v_1, \ldots, v_M\}$ represents the pursuer agents, and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ encodes the available communication links. A directed edge $(v_j, v_i) \in \mathcal{E}$ indicates that pursuer i can receive information from pursuer j, associated with a positive weight $\alpha_{ij} > 0$ if communication exists, and $\alpha_{ij} = 0$ otherwise. The adjacency matrix is defined as $\mathcal{A} = [\alpha_{ij}] \in \mathbb{R}^{M\times M}$, where $\alpha_{ii} = 0$ for all i. The in-degree matrix is given by $\mathcal{D} = \mathrm{diag}\{d_i\}$ with $d_i = \sum_{j=1}^{M} \alpha_{ij}$, and the corresponding Laplacian matrix is $\mathcal{L} = \mathcal{D} - \mathcal{A}$. To incorporate direct sensing of the evader by some pursuers, an augmented graph is constructed by adding a virtual node v_0 representing the evader. If pursuer i can directly detect the evader, a

directed edge (v_0, v_i) is added, with a corresponding pinning weight $\gamma_i > 0$. The influence of the evader is captured by the pinning matrix $\mathcal{R} = \mathrm{diag}\{\gamma_i\} \in \mathbb{R}^{M\times M}$.

In this study, the augmented graph is assumed to contain a spanning tree rooted at v_0, ensuring that the evader's state can be cooperatively estimated by all pursuers. Furthermore, communication links may be unreliable or faulty, leading to time-varying and uncertain edge weights, which are explicitly modeled in the following subsection.

2.2 Dynamics description

The dynamics of each pursuer agent i $(i = 1, \ldots, M)$ are described by a general nonlinear control-affine system:

$$\dot{x}_{p_i}(t) = f(x_{p_i}(t)) + g(x_{p_i}(t))u_{p_i}(t), \tag{1}$$

where $x_{p_i}(t) \in \mathbb{R}^n$ and $u_{p_i}(t) \in \mathbb{R}^m$ denote the state and control input of pursuer i, respectively. The function $f(\cdot) \in \mathbb{R}^n$ is a continuous nonlinear drift term, and $g(\cdot) \in \mathbb{R}^{n\times m}$ is a continuous control gain function.

The evader's dynamics follow a similar structure:

$$\dot{x}_e(t) = f(x_e(t)) + g(x_e(t))u_e(t), \tag{2}$$

where $x_e(t) \in \mathbb{R}^n$ and $u_e(t) \in \mathbb{R}^m$ denote the state and control input of the evader, respectively. Unless otherwise stated, the time argument t is omitted for brevity.

Assumption 1. The nonlinear function $f(\cdot)$ is locally Lipschitz continuous over a compact set $\Omega_x \subseteq \mathbb{R}^n$; that is, there exists a constant $\Pi_f > 0$, such that for any $x, \hat{x} \in \Omega_x$, $\|f(x) - f(\hat{x})\| \leqslant \Pi_f \|x - \hat{x}\|$. Additionally, the control gain function $g(\cdot)$ is uniformly bounded on Ω_x; that is, there exists a constant $\Pi_g > 0$ such that $\|g(x)\| \leqslant \Pi_g$, $\forall x \in \Omega_x$.

Assumption 2. The evader's control input $u_e(t)$ is uniformly bounded for all $t \geqslant 0$; that is, there exists a constant $u_m > 0$ such that $\|u_e(t)\| \leqslant u_m$, $\forall t \geqslant 0$. Furthermore, the evader's state trajectory remains within Ω_x, and it has full access to the global state information of all pursuers.

Remark 1. Assumptions 1 and 2 jointly ensure that the system dynamics remain well-posed within a compact domain. Specifically, Assumption 1 enforces local Lipschitz continuity and boundedness of $f(\cdot)$ and $g(\cdot)$, guaranteeing the existence and uniqueness of system trajectories. Assumption 2 restricts the evader's control input to be physically feasible, preventing divergence under realistic velocity and energy constraints. Although assuming full access to the pursuers' global state information may be idealized, it simplifies the game formulation and facilitates Nash-equilibrium analysis.

In the MPE game, each pursuer minimizes its distance to the evader, whereas the evader maximizes its separation from all pursuers. The evader is considered captured once any pursuer enters a specified capture radius, defined as follows.

Definition 1. The evader is considered captured if at least one pursuer's state trajectory enters the capture radius of the evader, that is,

$$\lim_{t\to\infty} \|x_{p_i}(t) - x_e(t)\| \leqslant \Delta_p, \quad \exists\, x_{p_i}(t) \in P, \tag{3}$$

where $P = \{x_{p_1}(t), \ldots, x_{p_M}(t)\}$ denotes the set of pursuers, and $\Delta_p > 0$ is a predefined capture radius.

To facilitate controller design, define the relative state between pursuer i and the evader as $z_i(t) = x_{p_i}(t) - x_e(t)$. The corresponding dynamics are given by

$$\dot{z}_i(t) = \mathcal{F}_i + \mathcal{G}_i u_{p_i}(t) + \mathcal{G}_e u_e(t),\ i = 1, \ldots, M, \tag{4}$$

where $\mathcal{F}_i = f(x_{p_i}) - f(x_e)$, $\mathcal{G}_i = g(x_{p_i})$, and $\mathcal{G}_e = -g(x_e)$. This establishes the differential game framework, in which each pursuer seeks to minimize $\|z_i(t)\|$, while the evader aims to maximize it.

2.3 Communication link faults

In real-world multi-agent systems, communication links are often affected by signal fading, hardware failures, environmental disturbances, or malicious attacks. Such factors lead to time-varying and uncertain network connectivity, which can severely degrade cooperative performance.

To model such effects, we define the faulty communication weights as

$$\begin{cases} \bar{\alpha}_{ij}(t) = \alpha_{ij} + \theta_{ij}^{\alpha}(t), \\ \bar{\gamma}_i(t) = \gamma_i + \theta_i^{\gamma}(t), \end{cases} \tag{5}$$

where $\bar{\alpha}_{ij}(t)$ and $\bar{\gamma}_i(t)$ denote the actual (possibly degraded) communication weights; $\theta_{ij}^{\alpha}(t)$, $\theta_i^{\gamma}(t)$ represent unknown time-varying fault signals affecting inter-pursuer and pursuer-to-evader communications, respectively.

This results in time-varying adjacency matrix $\mathcal{A}^{\theta}(t)$, pinning matrix $\mathcal{R}^{\theta}(t)$, and Laplacian matrix $\mathcal{L}^{\theta}(t)$, corresponding to the faulty communication topology. It should be noted that the adopted additive model mainly captures continuous degradations such as signal attenuation or bounded interference. However, it does not account for discontinuous phenomena like packet dropouts, intermittent link failures, or communication delays, which may induce abrupt connectivity changes and latency-related instability. Addressing these issues remains an important direction for future research.

Assumption 3. The communication fault signals $\theta_{ij}^{\alpha}(t)$ and $\theta_i^{\gamma}(t)$, along with their first derivatives, are bounded for all $t \geqslant 0$. Moreover, the signs of $\bar{\alpha}_{ij}(t)$ and $\bar{\gamma}_i(t)$ remain consistent with those of their nominal counterparts α_{ij} and γ_i, thereby preserving the original interaction structure of the network.

Lemma 1 ([17])**.** Let $\Gamma(t) = \mathcal{L}^{\theta}(t) + \mathcal{R}^{\theta}(t)$ be the time-varying Laplacian matrix of the augmented graph. Then, there exists a positive definite, time-varying diagonal matrix $\mathcal{Q}(t) = \mathrm{diag}\{q_i(t)\} \in \mathbb{R}^{M\times M}$ such that the following Lyapunov-like equation holds:

$$\Gamma^{\mathrm{T}}(t)\mathcal{Q}(t) + \mathcal{Q}(t)\Gamma(t) = \mathcal{P}(t), \tag{6}$$

where $\mathcal{P}(t)$ is a symmetric positive definite matrix. Moreover, both $\mathcal{Q}(t)$ and its time derivative $\dot{\mathcal{Q}}(t)$ are uniformly bounded for all $t \geqslant 0$.

Remark 2. Assumption 3 allows the communication fault model in (5) to characterize various real-world link degradations, such as signal attenuation, intermittent dropouts, and adversarial interference. By ensuring the boundedness of both the fault signals and their rates of change, the underlying graph structure maintains sufficient connectivity for distributed coordination. Lemma 1 further guarantees the existence of a time-varying positive definite matrix $\mathcal{Q}(t)$, which plays a critical role in constructing Lyapunov functions and facilitating stability analysis of the observer-based FTC framework. It is worth noting, however, that this assumption excludes adversarial scenarios where communication weights might flip their signs, effectively turning cooperative links into antagonistic ones. While this simplification facilitates stability analysis, extending the framework to address sign-varying weights through tools such as signed graph theory or resilient consensus methods constitutes an important avenue for future research.

Design objective. Given the nonlinear dynamics (1) and (2) and time-varying communication faults (5), the objective is to design a distributed FTC strategy that achieves the following goals.

(i) Each pursuer can accurately estimate the evader's state through a distributed observer, even under time-varying and uncertain communication link faults. That is, $\hat{x}_{p_{i0}} \to x_e$ as $t \to \infty$ for all $i = 1, \ldots, M$.

(ii) All pursuers capture the evader while ensuring that the derived strategies constitute a Nash equilibrium of the underlying PE game.

(iii) The closed-loop system remains stable, and all internal signals are uniformly ultimately bounded, even in the presence of communication faults and NN approximation errors.

These challenges motivate a hierarchical architecture combining a distributed adaptive observer and a single-critic NN-based control law. This structure enables real-time Nash strategy approximation while ensuring robustness to communication uncertainty. The overall structure is shown in Figure 1.

3 Main results

3.1 Distributed adaptive observer design

To mitigate the effects of communication link faults, each pursuer agent is equipped with a distributed adaptive observer to estimate the evader's state. Let $\hat{x}_{p_{i0}} \in \mathbb{R}^n$ denote the local estimate of the evader's state by the i-th pursuer, and define the global estimate vector as $\hat{x}_{p_0} = [\hat{x}_{p_{10}}^{\mathrm{T}}, \ldots, \hat{x}_{p_{M0}}^{\mathrm{T}}]^{\mathrm{T}}$. For each pursuer, a local observer error ζ_i is defined based on neighbor communications and direct evader observations (if available), defined as

$$\zeta_i = \sum_{j=1, j\neq i}^{M} \bar{\alpha}_{ij}(t)\left(\hat{x}_{p_{i0}} - \hat{x}_{p_{j0}}\right) + \bar{\gamma}_i(t)\left(\hat{x}_{p_{i0}} - x_e\right), \tag{7}$$

where $\bar{\alpha}_{ij}(t)$ and $\bar{\gamma}_i(t)$ are the corrupted communication weights defined in (5), which reflect time-varying faults in inter-pursuer and pursuer-to-evader communication, respectively.

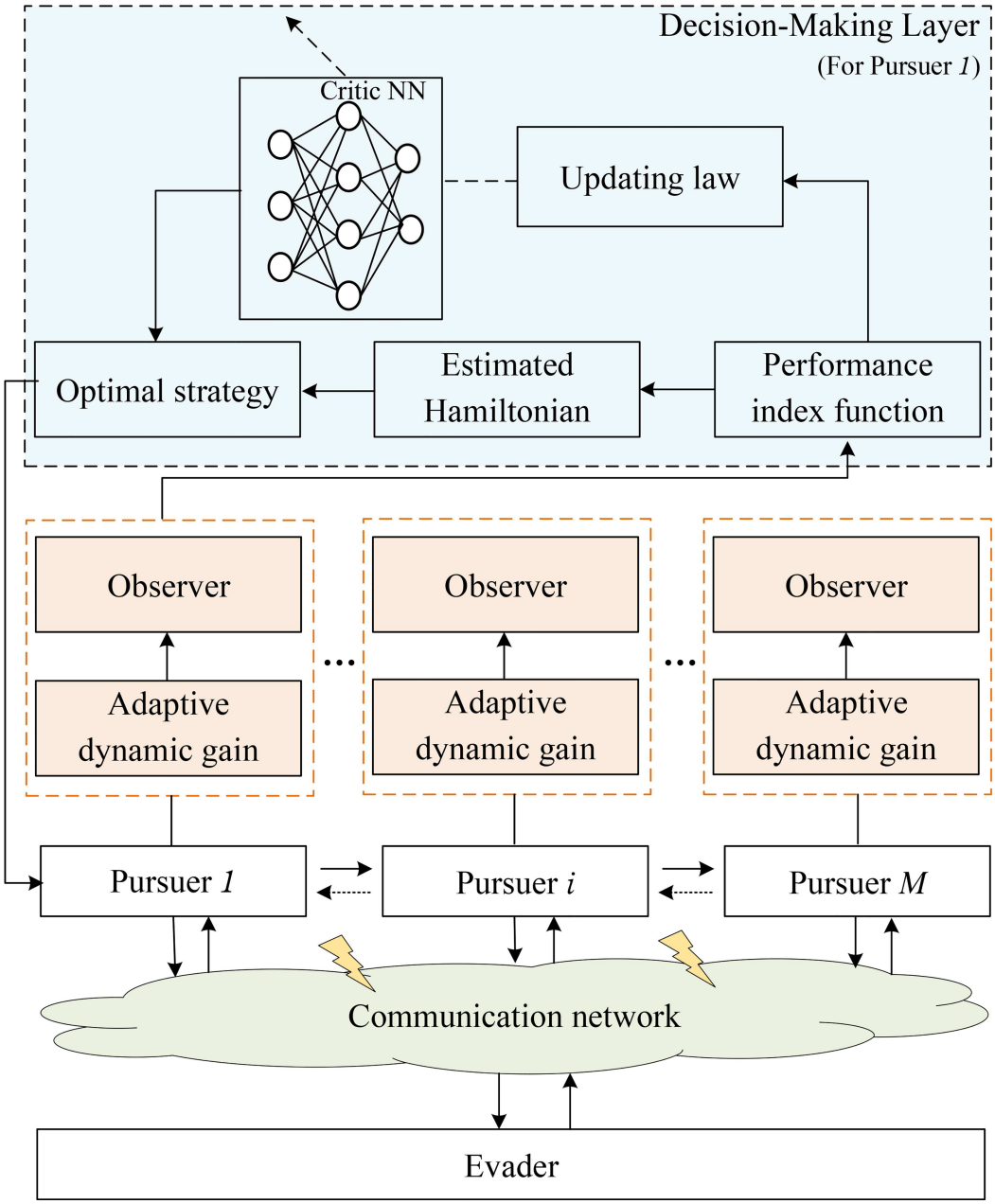

Figure 1 (Color online) Hierarchical control architecture for fault-tolerant MPE game under communication link faults.

The observer dynamics for the i-th pursuer are formulated as

$$\dot{\hat{x}}_{p_{i0}} = f(\hat{x}_{p_{i0}}) - \hbar \mathrm{sgn}(\zeta_i) - (\beta_i + k\dot{\beta}_i)\zeta_i, \tag{8}$$

where $\hbar$ is a design constant to be specified later, $\mathrm{sgn}(\zeta_i) \in \mathbb{R}_n$ denotes the element-wise sign function, and $k > 0$ is a positive constant that scales the dynamic gain component. The term β_i represents an adaptive observer gain, which evolves according to the following update law:

$$\dot{\beta}_i = \zeta_i^{\mathrm{T}} \zeta_i \tag{9}$$

with the initial condition $\beta_i(0) \geqslant k$, for all $i = 1, \ldots, M$.

Assumption 4. Each pursuer has access to its local error signal ζ_i in real time, which can be directly computed using onboard information and is utilized in both the observer update and adaptive gain law.

Remark 3. Although the observer error ζ_i includes unknown time-varying link weights $\bar{\alpha}_{ij}(t)$ and $\bar{\gamma}_i(t)$, it is not necessary to estimate these quantities explicitly. This is because the terms $\bar{\alpha}_{ij}(t)(\hat{x}_{p_{i0}} - \hat{x}_{p_{j0}})$ and $\bar{\gamma}_i(t)(\hat{x}_{p_{i0}} - x_e)$

can be directly obtained from available communication data. This structure allows for real-time implementation under uncertain link conditions and aligns with standard approaches in distributed estimation for systems with communication uncertainties, as discussed in [16].

Theorem 1. Consider the MPE system (1) and (2) with communication link faults as in (5). If the distributed observer (8) and adaptive gain law (9) are implemented and Assumptions 1–4 hold, then each pursuer's local estimate of the evader converges asymptotically to the true state. Specifically, $\lim_{t\to\infty}\|\hat{x}_{p_{i0}}(t)-x_e(t)\|=0,\forall i=1,\ldots,M$.

Proof. Let the estimation error of the i-th pursuer be $\tilde{x}_{p_{i0}}=\hat{x}_{p_{i0}}-x_e$, and define the global estimation error as $\tilde{x}_{p_0}=[\tilde{x}_{p_{10}}^{\mathrm{T}},\ldots,\tilde{x}_{p_{M0}}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{Mn}$. Similarly, denote the stacked observer estimates as $\hat{x}_{p_0}=[\hat{x}_{p_{10}}^{\mathrm{T}},\ldots,\hat{x}_{p_{M0}}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{Mn}$. From (7) and Lemma 1, the concatenated observer error can be compactly expressed as

$$\zeta=[\zeta_1^{\mathrm{T}},\ldots,\zeta_M^{\mathrm{T}}]^{\mathrm{T}}=(\Gamma(t)\otimes I_n)\tilde{x}_{p_0},\tag{10}$$

where $\Gamma(t)=\mathcal{L}^{\theta}(t)+\mathcal{R}^{\theta}(t)$ as defined in Lemma 1.

Differentiating ζ and substituting (1), (2), (4), and (8), we obtain the following error dynamics:

$$\begin{aligned}\dot{\zeta}&=(\Gamma(t)\otimes I_n)\dot{\tilde{x}}_{p_0}+(\dot{\Gamma}(t)\otimes I_n)\tilde{x}_{p_0}\\&=(\Gamma(t)\otimes I_n)\left(\dot{\hat{x}}_{p_0}-\dot{x}_e\otimes\mathbf{1}_M\right)+(\dot{\Gamma}(t)\otimes I_n)\tilde{x}_{p_0}\\&=(\Gamma(t)\otimes I_n)\left(\hat{\mathcal{F}}-\hbar\mathrm{sgn}(\zeta)-\bar{\mathcal{G}}_e(\mathbf{1}_M\otimes u_e)\right)-\left(\Gamma(t)(\mathcal{B}+k\dot{\mathcal{B}})\otimes I_n\right)\zeta+(\dot{\Gamma}(t)\otimes I_n)\tilde{x}_{p_0},\end{aligned}\tag{11}$$

where $\hat{\mathcal{F}}=[f(\hat{x}_{p_{10}})-f(x_e);\ldots;f(\hat{x}_{p_{M0}})-f(x_e)]$, $\mathrm{sgn}(\zeta)=[\mathrm{sgn}(\zeta_1),\ldots,\mathrm{sgn}(\zeta_M)]^{\mathrm{T}}\in\mathbb{R}^M$, $\bar{\mathcal{G}}_e=I_M\otimes g(x_e)$, $\mathcal{B}=\mathrm{diag}\{\beta_1,\ldots,\beta_M\}$ and $\dot{\mathcal{B}}=\mathrm{diag}\{\dot{\beta}_1,\ldots,\dot{\beta}_M\}$.

Next, consider the following Lyapunov candidate function:

$$V_\zeta=\frac{1}{2k}\sum_{i=1}^{M}\left(2\beta_i+k\dot{\beta}_i\right)q_i(t)\zeta_i^{\mathrm{T}}\zeta_i+\frac{1}{2}\sum_{i=1}^{M}(\beta_i-\beta_\zeta)^2,\tag{12}$$

where $\beta_\zeta>0$ is a design constant, and $q_i(t)$ are the time-varying positive weights satisfying the conditions in Lemma 1. Taking the time derivative of V_ζ and substituting the adaptive law (9) yields

$$\begin{aligned}\dot{V}_\zeta=&\frac{2}{k}\sum_{i=1}^{M}\left(k\dot{\beta}_i+\beta_i\right)q_i(t)\zeta_i^{\mathrm{T}}\dot{\zeta}_i+\frac{1}{k}\sum_{i=1}^{M}\dot{\beta}_iq_i(t)\zeta_i^{\mathrm{T}}\zeta_i+\frac{1}{2k}\sum_{i=1}^{M}\left(2\beta+k\dot{\beta}_i\right)\dot{q}_i(t)\zeta_i^{\mathrm{T}}\zeta_i+\sum_{i=1}^{M}(\beta_i-\beta_\zeta)\dot{\beta}_i\\=&\frac{2}{k}\zeta^{\mathrm{T}}\left[\left(k\dot{\mathcal{B}}+\mathcal{B}\right)\mathcal{Q}(t)\otimes I_n\right]\dot{\zeta}+\frac{1}{k}\zeta^{\mathrm{T}}\left(\dot{\mathcal{B}}\mathcal{Q}(t)\otimes I_n\right)\zeta\\&+\frac{1}{2k}\zeta^{\mathrm{T}}\left(\left(k\dot{\mathcal{B}}+2\mathcal{B}\right)\dot{\mathcal{Q}}(t)\otimes I_n\right)\zeta+\zeta^{\mathrm{T}}\left(\mathcal{B}\otimes I_n\right)\zeta-\beta_\zeta\zeta^{\mathrm{T}}\zeta.\end{aligned}\tag{13}$$

Substituting (11) into the first term of (13) yields

$$\begin{aligned}\frac{2}{k}\zeta^{\mathrm{T}}\left[\left(k\dot{\mathcal{B}}+\mathcal{B}\right)\mathcal{Q}(t)\otimes I_n\right]\dot{\zeta}=&\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\Gamma(t)\otimes I_n\right]\left(\hat{\mathcal{F}}-\hbar\mathrm{sgn}(\zeta)-\bar{\mathcal{G}}_e(\mathbf{1}_M\otimes u_e)\right)\\&-\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\Gamma(t)(\mathcal{B}+k\dot{\mathcal{B}})\otimes I_n\right]\zeta\\&+\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\dot{\Gamma}(t)\otimes I_n\right]\tilde{x}_{p_0}.\end{aligned}\tag{14}$$

By applying Assumptions 1 and 2, the following bounds hold $\|\hat{\mathcal{F}}\|\leqslant\Pi_f\|\tilde{x}_{p_0}\|$, let the design parameter satisfy $\hbar\geqslant\Pi_g\sqrt{M}\|u_m\|$. Moreover, from (10) and Lemma 1, $\|\tilde{x}_{p_0}\|\leqslant\|\zeta\|/\lambda_{\min}(\Gamma(t))$. Using Young's inequality to bound all cross terms, the following inequalities hold:

$$\begin{aligned}&\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\Gamma(t)\otimes I_n\right]\hat{\mathcal{F}}\leqslant\frac{\tau_0}{4k}\zeta^{\mathrm{T}}\left((k\dot{\mathcal{B}}+\mathcal{B})^2\otimes I_n\right)\zeta+\frac{\tau_1^2\Pi_f^2\zeta^{\mathrm{T}}\zeta}{k\tau_0\lambda_{\min}^2(\Gamma(t)\otimes I_n)},\\&\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\Gamma(t)\otimes I_n\right]\left(-\hbar\mathrm{sgn}(\zeta)-\bar{\mathcal{G}}_e(\mathbf{1}_M\otimes u_e)\right)\leqslant 0,\end{aligned}\tag{15}$$

where $\tau_0=\min_{\forall t\geqslant 0}\lambda_{\min}\{\mathcal{P}(t)\}$, $\tau_1=\max_{\forall t\geqslant 0}\lambda_{\max}\{\mathcal{P}(t)\}$. Here, $\lambda_{\min}\{\cdot\}$ and $\lambda_{\max}\{\cdot\}$ denote the minimum and maximum eigenvalues of the corresponding matrix, respectively.

Similarly, for the remaining terms, applying Lemma 1 and Young's inequality yields

$$-\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\Gamma(t)(\mathcal{B}+k\dot{\mathcal{B}})\otimes I_n\right]\zeta\leqslant-\frac{\tau_0}{k}\zeta^{\mathrm{T}}\left((k\dot{\mathcal{B}}+\mathcal{B})^2\otimes I_n\right)\zeta, \tag{16}$$

$$\frac{2}{k}\zeta^{\mathrm{T}}\left[(k\dot{\mathcal{B}}+\mathcal{B})\mathcal{Q}(t)\dot{\Gamma}(t)\otimes I_n\right]\tilde{x}_{p_0}\leqslant\frac{\tau_0}{4k}\zeta^{\mathrm{T}}\left[\left(k\dot{\mathcal{B}}+\mathcal{B}\right)^2\otimes I_n\right]\zeta+\frac{\tau_2^2\zeta^{\mathrm{T}}\zeta}{k\tau_0\lambda_{\min}^2(\Gamma(t)\otimes I_n)} \tag{17}$$

with $\tau_2=2\max_{\forall t\geqslant 0}\|\mathcal{Q}(t)\dot{\Gamma}(t)\|$. Then, substituting (16) and (17) into (13), we arrive at the following bound for the derivative of the Lyapunov function:

$$\begin{aligned}\dot{V}_\zeta\leqslant&-\frac{\tau_0}{2k}\zeta^{\mathrm{T}}\left((k\dot{\mathcal{B}}+\mathcal{B})^2\otimes I_n\right)\zeta+\frac{(\tau_1^2\Pi_f^2+\tau_2^2)\zeta^{\mathrm{T}}\zeta}{k\tau_0\lambda_{\min}^2(\Gamma(t)\otimes I_n)}+\frac{1}{k}\zeta^{\mathrm{T}}\left(\dot{\mathcal{B}}\mathcal{Q}(t)\otimes I_n\right)\zeta\\&+\frac{1}{2k}\zeta^{\mathrm{T}}\left(\left(k\dot{\mathcal{B}}+2\mathcal{B}\right)\dot{\mathcal{Q}}(t)\otimes I_n\right)\zeta+\zeta^{\mathrm{T}}\left(\mathcal{B}\otimes I_n\right)\zeta-\beta_\zeta\zeta^{\mathrm{T}}\zeta.\end{aligned} \tag{18}$$

The remaining terms can be bounded as follows:

$$\frac{1}{k}\zeta^{\mathrm{T}}\left(\dot{\mathcal{B}}\mathcal{Q}(t)\otimes I_n\right)\zeta\leqslant\frac{\tau_0}{8k}\zeta^{\mathrm{T}}\left(k^2\dot{\mathcal{B}}^2\otimes I_n\right)\zeta+\frac{2\bar{q}_m^2}{k^3\tau_0}\zeta^{\mathrm{T}}\zeta, \tag{19}$$

$$\frac{1}{2k}\zeta^{\mathrm{T}}\left(\left(k\dot{\mathcal{B}}+2\mathcal{B}\right)\dot{\mathcal{Q}}(t)\otimes I_n\right)\zeta\leqslant\frac{\tau_0}{8k}\zeta^{\mathrm{T}}\left(k^2\dot{\mathcal{B}}^2\otimes I_n\right)\zeta+\frac{\ddot{\bar{q}}_m^2}{2k\tau_0}\zeta^{\mathrm{T}}\zeta+\frac{\tau_0}{8k}\zeta^{\mathrm{T}}\left(\mathcal{B}^2\otimes I_n\right)\zeta+\frac{2\ddot{\bar{q}}_m^2}{k\tau_0}\zeta^{\mathrm{T}}\zeta, \tag{20}$$

$$\zeta^{\mathrm{T}}\left(\mathcal{B}\otimes I_n\right)\zeta\leqslant\frac{\tau_0}{8k}\zeta^{\mathrm{T}}\left(\mathcal{B}^2\otimes I_n\right)\zeta+\frac{2k}{\tau_0}\zeta^{\mathrm{T}}\zeta, \tag{21}$$

where $\bar{q}_m=\max_{\{\forall t\geqslant 0,i\in\mathcal{V}\}}\{q_i(t)\}$ and $\dot{\bar{q}}_m=\max_{\{\forall t\geqslant 0,i\in\mathcal{V}\}}\{\dot{q}_i(t)\}$.

Combining the above results, we derive the final bound

$$\dot{V}_\zeta\leqslant-\frac{\tau_0}{4k}\zeta^{\mathrm{T}}\left((k\dot{\mathcal{B}}+\mathcal{B})^2\otimes I_n\right)\zeta-\beta_\zeta\zeta^{\mathrm{T}}\zeta+\bar{\tau}\zeta^{\mathrm{T}}\zeta, \tag{22}$$

where $\bar{\tau}=\frac{2\tau_1^2\Pi_f^2+\tau_2^2}{k\tau_0\lambda_{\min}^2(\Gamma(t)\otimes I_n)}+\frac{2\bar{q}_m^2}{k^3\tau_0}+\frac{5\ddot{\bar{q}}_m^2}{2k\tau_0}+\frac{2k}{\tau_0}$. Thus, by choosing $\beta_\zeta\geqslant\bar{\tau}$, it follows that

$$\dot{V}_\zeta\leqslant-\frac{\tau_0}{4k}\zeta^{\mathrm{T}}\left((k\dot{\mathcal{B}}+\mathcal{B})^2\otimes I_n\right)\zeta, \tag{23}$$

which implies that $\zeta(t)\in L_2\bigcap L_\infty$, and all signals in the closed-loop system remain bounded. Finally, by invoking Barbalat's lemma, we conclude that $\lim_{t\to\infty}\|\hat{x}_{p_{i0}}-x_e\|=0$. This completes the proof.

Remark 4. The time-varying communication link faults hinder accurate and timely command transmission in MPE games. To address this, a novel distributed adaptive observer is proposed with the following features. (i) Fully distributed operation that fuses relative estimates among neighboring pursuers with direct evader measurements (when available), improving robustness against intermittent or faulty links. (ii) Each observer includes an adaptive gain β_i that evolves with the local error energy $\zeta_i^{\mathrm{T}}\zeta_i$, allowing dynamic adjustment of the convergence speed. The additional term $(\beta_i+k\dot{\beta}_i)$ accelerates adaptation and enhances stability under uncertainties. (iii) A feedback term $\mathrm{sgn}(\zeta_i)$ and implemented as $\tanh(\zeta_i/\varepsilon)$ with a small $\varepsilon>0$ that suppresses chattering, improves numerical stability, and enables real-time implementation.

3.2 Solution of MPE game

Consider the local error between the i-th pursuer and the estimated evader state, denoted as $\bar{z}_i=x_{p_i}-\hat{x}_{p_{i0}}$. Additionally, define the aggregate error from the evader to all pursuers as $z_0=\sum_i^M(x_e-x_{p_i})$. The performance index for the i-th pursuer is then defined as

$$\mathcal{J}_i(\bar{z}_i,u_{p_i})=\int_0^\infty\left[\bar{z}_i^{\mathrm{T}}Q_i\bar{z}_i+u_{p_i}^{\mathrm{T}}R_iu_{p_i}\right]\mathrm{d}t, \tag{24}$$

where Q_i and R_i are symmetric positive-definite matrices.

Similarly, the performance index for the evader is given by

$$\mathcal{J}_0(z_0,u_e)=\int_0^\infty\left[z_0^{\mathrm{T}}Q_0z_0+u_e^{\mathrm{T}}R_0u_e\right]\mathrm{d}t, \tag{25}$$

where Q_0 is a symmetric negative-definite matrix, reflecting that the objective of the evader is contrary to that of the pursuers. Additionally, R_0 is a symmetric positive-definite matrix.

For the i-th pursuer and the evader, when both adopt their respective optimal strategies, the following expressions hold:

$$\mathcal{J}_i^* = \min_{u_{p_i}} \int_0^\infty \left[\bar{z}_i^{\mathrm{T}} Q_i \bar{z}_i + u_{p_i}^{\mathrm{T}} R_i u_{p_i}\right] \mathrm{d}t, \tag{26}$$

$$\mathcal{J}_0^* = \min_{u_e} \int_0^\infty \left[z_0^{\mathrm{T}} Q_0 z_0 + u_e^{\mathrm{T}} R_0 u_e\right] \mathrm{d}t. \tag{27}$$

In this context, each pursuer aims to minimize its distance to the observed evader state in (26) despite communication link faults, while the evader selects its optimal control according to (27). These interactions naturally define a Nash equilibrium for the PE game.

Definition 2. The control strategies $u_{p_1}^*, \ldots, u_{p_M}^*$ for the pursuers and u_e^* for the evader constitute a Nash equilibrium if the following conditions are met:

$$\mathcal{J}_i(u_{p_1}^*, \ldots, u_{p_M}^*, u_e^*) \leqslant \mathcal{J}_i(u_{p_1}^*, \ldots, u_{p_i}, \ldots, u_{p_M}^*, u_e^*), \tag{28}$$

$$\mathcal{J}_0(u_e^*, u_{p_1}^*, \ldots, u_{p_M}^*) \leqslant \mathcal{J}_0(u_e, u_{p_1}^*, \ldots, u_{p_M}^*). \tag{29}$$

Based on (8) and (24), the Hamiltonian function for the i-th pursuer is defined as follows:

$$\mathcal{H}_i = \bar{z}_i^{\mathrm{T}} Q_i \bar{z}_i + u_{p_i}^{\mathrm{T}} R_i u_{p_i} + (\nabla \mathcal{J}_i)^{\mathrm{T}} \big(\bar{\mathcal{F}}_i + \mathcal{G}_i u_{p_i}\big), \tag{30}$$

where $\nabla \mathcal{J}_i$ is the partial derivatives of $\bar{z}_i$ and $\bar{\mathcal{F}}_i = f(x_{p_i}) - f(\hat{x}_{p_{i0}}) - \hbar \mathrm{sgn}(\zeta_i) - (\beta_i + k\dot{\beta}_i)\zeta_i$. Similarly, by (25) and the Hamiltonian function for the evader is given by

$$\mathcal{H}_0 = z_0^{\mathrm{T}} Q_0 z_0 + u_e^{\mathrm{T}} R_0 u_e - (\nabla \mathcal{J}_0)^{\mathrm{T}} \sum_{i=1}^{M} \big(\mathcal{F}_i + \mathcal{G}_i u_{p_i} + \mathcal{G}_e u_e\big), \tag{31}$$

where $\nabla \mathcal{J}_0$ are the partial derivatives of z_0.

Leveraging the stationary conditions, it logically follows that:

$$\frac{\partial \mathcal{H}_i}{\partial u_{p_i}} = 0 \Rightarrow u_{p_i}^* = -\frac{1}{2} R_i^{-1} \mathcal{G}_i^{\mathrm{T}} \nabla \mathcal{J}_i^*, \tag{32}$$

$$\frac{\partial \mathcal{H}_0}{\partial u_e} = 0 \Rightarrow u_e^* = \frac{M}{2} R_0^{-1} \mathcal{G}_e^{\mathrm{T}} \nabla \mathcal{J}_0^*. \tag{33}$$

Substituting the optimal strategies (32) and (33) into (30) and (31) yields the corresponding HJI equations:

$$\bar{z}_i^{\mathrm{T}} Q_i \bar{z}_i + \nabla \mathcal{J}_i^{*\mathrm{T}} \bar{\mathcal{F}}_i - \frac{1}{4} \nabla \mathcal{J}_i^{*\mathrm{T}} \mathcal{G}_i R_i^{-1} \mathcal{G}_i{}^{\mathrm{T}} \nabla \mathcal{J}_i^* = 0, \tag{34}$$

$$z_0^{\mathrm{T}} Q_0 z_0 - \sum_{i=1}^{M} \nabla \mathcal{J}_0^{*\mathrm{T}} \mathcal{F}_i - \frac{M^2}{4} \nabla \mathcal{J}_0^{*\mathrm{T}} \mathcal{G}_e R_0^{-1} \mathcal{G}_e^{\mathrm{T}} \nabla \mathcal{J}_0^* + \sum_{i=1}^{M} \frac{1}{2} \nabla \mathcal{J}_0^{*\mathrm{T}} \mathcal{G}_i R_i^{-1} \mathcal{G}_i{}^{\mathrm{T}} \nabla \mathcal{J}_i^* = 0. \tag{35}$$

The capturability of the pursuers and the properties of the Nash equilibrium are further analyzed in Theorem 2.

Theorem 2. For the MPE system (1) and (2) under the distributed adaptive observer (8) with adaptive law (9), and employing the optimal strategies in (32) and (33), the evader is guaranteed to be captured. Moreover, these strategies form a Nash equilibrium for the MPE game.

Proof. The Lyapunov function is selected as $V_i = \mathcal{J}_i^*$ to prove the capturability between the multi-pursuers and the evader. Integrating this definition with the optimal strategies outlined in (32) and (33), yields $V_i > 0$. Furthermore, the time derivative of V_i can be expressed as follows:

$$\dot{V}_i = \nabla \mathcal{J}_i^{*\mathrm{T}} \big(\bar{\mathcal{F}}_i + \mathcal{G}_i u_{p_i}\big) = \nabla \mathcal{J}_i^{*\mathrm{T}} \bar{\mathcal{F}}_i - \frac{1}{2} \nabla \mathcal{J}_i^{*\mathrm{T}} \mathcal{G}_i R_i^{-1} \mathcal{G}_i{}^{\mathrm{T}} \nabla \mathcal{J}_i^*. \tag{36}$$

By further incorporating the HJI equations from (34), the above expression can be simplified as follows:

$$\dot{V}_i = -\bar{z}_i^{\mathrm{T}} Q_i \bar{z}_i - \frac{1}{4} \nabla \mathcal{J}_i^{*\mathrm{T}} \mathcal{G}_i R_i^{-1} \mathcal{G}_i{}^{\mathrm{T}} \nabla \mathcal{J}_i^* \leqslant -\lambda_{\min}(Q_i) \|\bar{z}_i\|^2 - \frac{1}{4} \lambda_{\min}(\mathcal{D}_i) \|\nabla \mathcal{J}_i^*\|^2 \leqslant 0. \tag{37}$$

This analysis demonstrates that the optimized strategy in (32) guarantees asymptotic convergence of the local error $\|\bar{z}_i\| \to 0$ as $t \to \infty$. Specifically, it guarantees that $\|x_{p_i} - \hat{x}_{p_i0}\| \to 0$ which according to Theorem 1 is equivalent to $\|x_{p_i} - x_e\| \to 0$. Moreover, this result establishes the capturability of the evader under the proposed FTC framework.

Next, the performance index $\mathcal{J}_0$ is rewritten as shown below to establish the Nash equilibrium.

$$\begin{aligned}\mathcal{J}_0 &= \int_0^\infty \left[z_0^{\mathrm{T}} Q_0 z_0 + u_e^{\mathrm{T}} R_0 u_e\right] \mathrm{d}t + \int_0^\infty \dot{V}_0 \, \mathrm{d}t + V_0(z_0(0)) \\ &= \int_0^\infty \left[z_0^{\mathrm{T}} Q_0 z_0 + u_e^{\mathrm{T}} R_0 u_e\right] \mathrm{d}t + V_0(z_0(0)) - \int_0^\infty \left[\nabla \mathcal{J}_0^{*\mathrm{T}} \sum_{i=1}^{M} \left(\mathcal{F}_i + \mathcal{G}_i u_{p_i} + \mathcal{G}_e u_e\right)\right] \mathrm{d}t. \end{aligned} \tag{38}$$

Here, $V_0 = \mathcal{J}_0^*$, and $V_0(z_0(\infty)) = V_0(0) = 0$ is defined as a boundary condition, ensuring the convergence of the Lyapunov function. Using the definition of (35), which can be expanded as

$$\sum_{i=1}^{M} \nabla \mathcal{J}_0^{*\mathrm{T}} \mathcal{F}_i = z_0^{\mathrm{T}} Q_0 z_0 + u_e^{*\mathrm{T}} R_0 u_e^* - \nabla \mathcal{J}_0^{*\mathrm{T}} \left(\mathcal{G}_i u_{p_i}^* + \mathcal{G}_e u_e^*\right). \tag{39}$$

The performance index $\mathcal{J}_0$ can be further rewritten as follows:

$$\begin{aligned}\mathcal{J}_0 =& \int_0^\infty \left[u_e^{\mathrm{T}} R_0 u_e + u_e^{*\mathrm{T}} R_0 u_e^*\right] \mathrm{d}t + \int_0^\infty \Bigg[- 2u_e^{*\mathrm{T}} R_0 u_e^* + \nabla \mathcal{J}_0^{*\mathrm{T}} \\ & \times \sum_{i=1}^{M} \left(\mathcal{G}_i \left(u_{p_i} - u_{p_i}^*\right) + \mathcal{G}_e \left(u_e - u_e^*\right)\right) \Bigg] \mathrm{d}t + V_0(z_0(0)) \\ =& \int_0^\infty \left[(u_e - u_e^*)^{\mathrm{T}} R_0 (u_e - u_e^*)\right] \mathrm{d}t + V_0(z_0(0)) \\ & + \int_0^\infty \left[2u_e^{*\mathrm{T}} R_0 \left(u_e - u_e^*\right) + \nabla \mathcal{J}_0^{*\mathrm{T}} \sum_{i=1}^{M} \left(\mathcal{G}_i \left(u_{p_i} - u_{p_i}^*\right) + \mathcal{G}_e \left(u_e - u_e^*\right)\right)\right] \mathrm{d}t. \end{aligned} \tag{40}$$

By rearranging terms and substituting $u_{p_i} = u_{p_i}^*$, Eq. (40) simplifies to

$$\mathcal{J}_0 = \int_0^\infty (u_e - u_e^*)^{\mathrm{T}} R_0 (u_e - u_e^*) \, \mathrm{d}t + V_0(z_0(0)). \tag{41}$$

From this, it follows that

$$\mathcal{J}_0(u_e^*, u_{p_1}^*, \ldots, u_{p_M}^*) \leqslant \mathcal{J}_0(u_e, u_{p_1}^*, \ldots, u_{p_M}^*), \tag{42}$$

indicating that the strategy u_e^* minimizes the performance index $\mathcal{J}_0$ of the evader when the strategies $u_{p_1}^*, \ldots, u_{p_M}^*$ of the pursuers are fixed.

Similarly, the performance index $\mathcal{J}_i$ for the evader can be expressed as

$$\mathcal{J}_i = \int_0^\infty \left[\left(u_{p_i} - u_{p_i}^*\right)^{\mathrm{T}} R_i \left(u_{p_i} - u_{p_i}^*\right)\right] \mathrm{d}t + V_i(\bar{z}_i(0)). \tag{43}$$

Thus, we have

$$\mathcal{J}_i(u_{p_1}^*, \ldots, u_{p_M}^*, u_e^*) \leqslant \mathcal{J}_i(u_{p_1}^*, \ldots, u_{p_i}, \ldots, u_{p_M}^*, u_e^*), \tag{44}$$

indicating that the strategy $u_{p_i}^*$ minimizes the performance index $\mathcal{J}_i$ when the strategy of the evader u_e^* is fixed.

The strategies $u_{p_1}^*, \ldots, u_{p_M}^*$ of the pursuers and the strategy u_e^* of the evader independently minimize their respective performance indexes. Thus, the combination $(u_{p_i}^*, u_e^*)$ constitutes a Nash equilibrium for the game.

3.3 Online critic NN implementation

The critic NN leverages its approximation capability to solve the HJI equations in (34) and (35), approximating the performance index functions $\mathcal{J}_i^*$ and $\mathcal{J}_0^*$ in (26) and (27), as follows:

$$\mathcal{J}_i^*(\bar{z}_i) = \mathcal{W}_{c_i}^{\mathrm{T}} \phi_{c_i}(\bar{z}_i) + \iota_{c_i}(\bar{z}_i), \quad i = 1, \ldots, M, \tag{45a}$$

$$\mathcal{J}_0^*(z_0) = \mathcal{W}_{c_0}^{\mathrm{T}}\phi_{c_0}(z_0) + \iota_{c_0}(z_0) \tag{45b}$$

together with their partial derivatives

$$\nabla\mathcal{J}_i^*(\bar{z}_i) = \nabla\phi_{c_i}^{\mathrm{T}}(\bar{z}_i)\mathcal{W}_{c_i} + \nabla\iota_{c_i}(\bar{z}_i), \quad i = 1, \ldots, M, \tag{46a}$$

$$\nabla\mathcal{J}_0^*(z_0) = \nabla\phi_{c_0}^{\mathrm{T}}(z_0)\mathcal{W}_{c_0} + \nabla\iota_{c_0}(z_0), \tag{46b}$$

where $\mathcal{W}_{c_i} \in R^r$ and $\mathcal{W}_{c_0} \in R^r$ are the ideal weight vectors, $\phi_{c_i}(\bar{z}_i)$ and $\phi_{c_0}(z_0)$ are the activation functions, $\iota_{c_i}(\bar{z}_i)$ and $\iota_{c_0}(z_0)$ are the approximation errors, and r is the number of neurons in the hidden layer. As r increases, the approximation errors $\iota_{c_i}(\bar{z}_i)$ and $\iota_{c_0}(z_0)$ converge to zero. Moreover, $\nabla\phi_{c_i}(\bar{z}_i)$, $\nabla\phi_{c_0}(z_0)$, $\nabla\iota_{c_i}(\bar{z}_i)$ and $\nabla\iota_{c_0}(z_0)$ represent their corresponding partial derivatives.

Assumption 5. The ideal weights $\mathcal{W}_{c_i}$ and $\mathcal{W}_{c_0}$ are bounded by positive constants $\bar{\mathcal{W}}_{c_i}$ and $\bar{\mathcal{W}}_{c_0}$, respectively. Therefore, $\|\mathcal{W}_{c_i}\| \leqslant \bar{\mathcal{W}}_{c_i}$ and $\|\mathcal{W}_{c_0}\| \leqslant \bar{\mathcal{W}}_{c_0}$.

Assumption 6. The activation functions ϕ_{c_i} and ϕ_{c_0}, along with their partial derivatives $\nabla\phi_{c_i}$ and $\nabla\phi_{c_0}$ are bounded by positive constants $\bar{\phi}_{c_i}$, $\bar{\phi}_{dc_i}$, $\bar{\phi}_{c_0}$, and $\bar{\phi}_{dc_0}$, respectively. Specifically, $\|\phi_{c_i}\| \leqslant \bar{\phi}_{c_i}$, $\|\nabla\phi_{c_i}\| \leqslant \bar{\phi}_{dc_i}$, $\|\phi_{c_0}\| \leqslant \bar{\phi}_{c_0}$, $\|\nabla\phi_{c_0}\| \leqslant \bar{\phi}_{dc_0}$.

Assumption 7. The approximate residuals ι_{c_i} and ι_{c_0} as well as their partial derivatives $\nabla\iota_{c_i}$ and $\nabla\iota_{c_0}$ are bounded by positive constants $\bar{\iota}_{c_i}$, $\bar{\iota}_{dc_i}$, $\bar{\iota}_{c_0}$, and $\bar{\iota}_{dc_0}$, respectively. Consequently, $\|\iota_{c_i}\| \leqslant \bar{\iota}_{c_i}$, $\|\nabla\iota_{c_i}\| \leqslant \bar{\iota}_{dc_i}$, $\|\iota_{c_0}\| \leqslant \bar{\iota}_{c_0}$, $\|\nabla\iota_{c_0}\| \leqslant \bar{\iota}_{dc_0}$.

Based on these assumptions, the optimal strategies from (32) and (33) can be represented using NN-based approximations as follows:

$$u_{p_i}^* = -\frac{1}{2}R_i^{-1}\mathcal{G}_i^{\mathrm{T}}\Big(\nabla\phi_{c_i}^{\mathrm{T}}(\bar{z}_i)\mathcal{W}_{c_i} + \nabla\iota_{c_i}(\bar{z}_i)\Big), \tag{47a}$$

$$u_e^* = \frac{M}{2}R_0^{-1}\mathcal{G}_e^{\mathrm{T}}\Big(\nabla\phi_{c_0}^{\mathrm{T}}(z_0)\mathcal{W}_{c_0} + \nabla\iota_{c_0}(z_0)\Big). \tag{47b}$$

Substituting these optimal strategies into (34) and (35) yields

$$\bar{z}_i^{\mathrm{T}}Q_i\bar{z}_i + \mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\bar{\mathcal{F}}_i - \frac{1}{4}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i} = \sigma_{c_i}, \tag{48a}$$

$$z_0^{\mathrm{T}}Q_0z_0 - \sum_{i=1}^{M}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{F}_i - \frac{M^2}{4}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0} + \frac{1}{2}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\sum_{i=1}^{M}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i} = \sigma_{c_0} \tag{48b}$$

along with the expressions for the approximation residual errors, which are given as follows:

$$\sigma_{c_i} = -\nabla\iota_{c_i}^{\mathrm{T}}\left(\bar{\mathcal{F}}_i - \frac{1}{2}\mathcal{D}_i\left(\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i} + \nabla\iota_{c_i}\right)\right) - \frac{1}{4}\nabla\iota_{c_i}^{\mathrm{T}}\mathcal{D}_i\nabla\iota_{c_i}, \tag{49a}$$

$$\begin{aligned}\sigma_{c_0} = {}& \nabla\iota_{c_0}^{\mathrm{T}}\sum_{i=1}^{M}\left(\mathcal{F}_i - \frac{1}{2}\mathcal{D}_i\left(\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i} + \nabla\iota_{c_i}\right) + \frac{M}{2}\mathcal{D}_0\left(\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0} + \nabla\iota_{c_0}\right)\right)\\ & - \frac{M^2}{4}\nabla\iota_{c_0}^{\mathrm{T}}\mathcal{D}_0\nabla\iota_{c_0} - \frac{1}{2}\sum_{i=1}^{M}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\iota_{c_i}.\end{aligned} \tag{49b}$$

Since the ideal weight vectors $\mathcal{W}_{c_i}$ and $\mathcal{W}_{c_0}$ are unknown, they are replaced by their estimates $\hat{\mathcal{W}}_{c_i}$ and $\hat{\mathcal{W}}_{c_0}$, yielding the following performance index approximations:

$$\hat{\mathcal{J}}_i(\bar{z}_i) = \hat{\mathcal{W}}_{c_i}^{\mathrm{T}}\phi_{c_i}(\bar{z}_i), \tag{50a}$$

$$\hat{\mathcal{J}}_0(z_0) = \hat{\mathcal{W}}_{c_0}^{\mathrm{T}}\phi_{c_0}(z_0). \tag{50b}$$

The corresponding partial derivatives can be inferred from these approximations

$$\nabla\hat{\mathcal{J}}_i(\bar{z}_i) = \nabla\phi_{c_i}^{\mathrm{T}}(\bar{z}_i)\hat{\mathcal{W}}_{c_i}, \tag{51a}$$

$$\nabla\hat{\mathcal{J}}_0(z_0) = \nabla\phi_{c_0}^{\mathrm{T}}(z_0)\hat{\mathcal{W}}_{c_0}. \tag{51b}$$

Then, the optimal strategies in (47a) and (47b) can be expressed as follows:

$$\hat{u}_{p_i} = -\frac{1}{2}R_i^{-1}\mathcal{G}_i^{\mathrm{T}}\nabla\phi_{c_i}^{\mathrm{T}}(\bar{z}_i)\hat{\mathcal{W}}_{c_i}, \tag{52a}$$

$$\hat{u}_e = \frac{M}{2}R_0^{-1}\mathcal{G}_e^{\mathrm{T}}\nabla\phi_{c_0}^{\mathrm{T}}(z_0)\hat{\mathcal{W}}_{c_0}. \tag{52b}$$

The estimated weight vectors are used to express the dynamics of $\bar{z}_i$ and z_0 as follows:

$$\dot{\bar{z}}_i = \bar{\mathcal{F}}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}, \tag{53a}$$

$$\dot{z}_0 = -\sum_{i=1}^{M}\left(\mathcal{F}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i} + \frac{M}{2}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right). \tag{53b}$$

Correspondingly, the Hamiltonian functions can be expressed in terms of the estimated parameters as

$$\hat{\mathcal{H}}_i = \bar{z}_i^{\mathrm{T}}Q_i\bar{z}_i + \hat{\mathcal{W}}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\bar{\mathcal{F}}_i - \frac{1}{4}\hat{\mathcal{W}}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i} \triangleq \varepsilon_{c_i}, \tag{54a}$$

$$\hat{\mathcal{H}}_0 = z_0^{\mathrm{T}}Q_0z_0 - \sum_{i=1}^{M}\hat{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{F}_i - \frac{M^2}{4}\hat{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0} + \sum_{i=1}^{M}\frac{1}{2}\hat{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i} \triangleq \varepsilon_{c_0}, \tag{54b}$$

where ε_{c_i} and ε_{c_0} denote the residual errors resulting from replacing the ideal weights with their estimates.

The gradient descent method is employed to minimize the squared residual error

$$E = \frac{1}{2}\sum_{i=1}^{M}\varepsilon_{c_i}^{\mathrm{T}}\varepsilon_{c_i} + \frac{1}{2}\varepsilon_{c_0}^{\mathrm{T}}\varepsilon_{c_0}. \tag{55}$$

To ensure stability during optimization, the weight vectors $\hat{\mathcal{W}}_{c_i}$ and $\hat{\mathcal{W}}_{c_0}$ are updated as follows:

$$\dot{\hat{\mathcal{W}}}_{c_i} = -l_i\frac{\bar{\kappa}_i}{\hbar_i}\varepsilon_{c_i} + \frac{l_i}{4}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\hat{\mathcal{W}}_{c_i} - l_i\Big(\Upsilon_{i1}\hat{\mathcal{W}}_{c_i} - \Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\Big), \tag{56a}$$

$$\dot{\hat{\mathcal{W}}}_{c_0} = -l_0\frac{\bar{\kappa}_0}{\hbar_0}\varepsilon_{c_0} + l_0\frac{M^2}{4}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\hat{\mathcal{W}}_{c_0} - l_0\Big(\Upsilon_{01}\hat{\mathcal{W}}_{c_0} - \Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\Big), \tag{56b}$$

where the auxiliary variables are defined as shown below:

$$\kappa_i = \nabla\phi_{c_i}\left(\bar{\mathcal{F}}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\right),\ \kappa_0 = -\nabla\phi_{c_0}\sum_{i=1}^{M}\left(\mathcal{F}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i} + \frac{M}{2}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right),$$
$$\bar{\kappa}_i = \frac{\kappa_i}{1+\kappa_i^{\mathrm{T}}\kappa_i},\ \bar{\kappa}_0 = \frac{\kappa_0}{1+\kappa_0^{\mathrm{T}}\kappa_0},\ \hbar_i = 1+\kappa_i^{\mathrm{T}}\kappa_i,\ \hbar_0 = 1+\kappa_0^{\mathrm{T}}\kappa_0.$$

Here, $l_i > 0$ and $l_0 > 0$ are the critic NN learning rates, balancing convergence speed and stability, while Υ_{01}, Υ_{02}, Υ_{i1}, and Υ_{i2} are tuning parameters ensuring effective and robust optimization.

By combining (48a) and (48b), (54a) and (54b), the residual errors ε_{c_i} and ε_{c_0} can be expressed as

$$\varepsilon_{c_i} = \sigma_{c_i} - \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\kappa_i + \frac{1}{4}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}, \tag{57a}$$

$$\varepsilon_{c_0} = \sigma_{c_0} + \tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\kappa_0 + \frac{M^2}{4}\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0} - \frac{1}{2}\sum_{i=1}^{M}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}. \tag{57b}$$

For analytical convenience, define the weight estimation errors as $\tilde{\mathcal{W}}_{c_i} = \mathcal{W}_{c_i} - \hat{\mathcal{W}}_{c_i}$ and $\tilde{\mathcal{W}}_{c_0} = \mathcal{W}_{c_0} - \hat{\mathcal{W}}_{c_0}$. Substituting these into (56a) and (56b) yields the following error dynamics:

$$\begin{aligned}\dot{\tilde{\mathcal{W}}}_{c_i} =& -l_i\bar{\kappa}_i\bar{\kappa}_i^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i} + l_i\frac{\bar{\kappa}_i}{\hbar_i}\sigma_{c_i} + \frac{l_i}{4}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\tilde{\mathcal{W}}_{c_i}\\ &-\frac{l_i}{4}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\hat{\mathcal{W}}_{c_i} + l_i\Big(\Upsilon_{i1}\hat{\mathcal{W}}_{c_i} - \Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\Big),\end{aligned} \tag{58a}$$

$$\dot{\tilde{\mathcal{W}}}_{c_0} = l_0\bar{\kappa}_0\bar{\kappa}_0^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0} + l_0\frac{\bar{\kappa}_0}{\hbar_0}\sigma_{c_0} + \frac{l_0M^2}{4}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\tilde{\mathcal{W}}_{c_0} - \frac{l_0}{2}\sum_{i=1}^{M}\nabla\phi_{c_0}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\tilde{\mathcal{W}}_{c_i}$$
$$- \frac{l_0M^2}{4}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\hat{\mathcal{W}}_{c_0} + l_0\left(\Upsilon_{01}\hat{\mathcal{W}}_{c_0} - \Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right). \tag{58b}$$

These dynamics characterize the evolution of $\tilde{\mathcal{W}}_{c_i}$ and $\tilde{\mathcal{W}}_{c_0}$, which are crucial in analyzing the convergence and stability of the learning system.

Theorem 3. Consider the MPE game (1) and (2) with performance index functions (45a) and (45b), the distributed adaptive observer (8) and (9), optimal strategies in (47a) and (47b), and weight tuning laws (56a) and (56b) under Assumptions 5–7. Then, the relative states $\bar{z}_i$, z_0 and the weight estimation errors $\tilde{\mathcal{W}}_{c_i}$, $\tilde{\mathcal{W}}_{c_0}$ are uniformly ultimately bounded.

Proof. Take the total Lyapunov function

$$V_{\mathrm{HJI}} = \sum_{i=1}^{M}\left(V_i + \frac{1}{2}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}l_i^{-1}\tilde{\mathcal{W}}_{c_i}\right) + V_0 + \frac{1}{2}\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}l_0^{-1}\tilde{\mathcal{W}}_{c_0}, \tag{59}$$

where $V_i = \mathcal{J}_i^*$ and $V_0 = \mathcal{J}_0^*$ under the approximated optimal solutions in (52a) and (52b), respectively.

The time derivative of V_{HJI} is indicated as follows:

$$\dot{V}_{\mathrm{HJI}} = \sum_{i=1}^{M}\left(\nabla\mathcal{J}_i^{*\mathrm{T}}\left(\bar{\mathcal{F}}_i + \mathcal{G}_i\hat{u}_{p_i}\right) + \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}l_i^{-1}\dot{\tilde{\mathcal{W}}}_{c_i}\right) - \nabla\mathcal{J}_0^{*\mathrm{T}}\sum_{i=1}^{M}\left(\mathcal{F}_i + \mathcal{G}_i\hat{u}_{p_i} + \mathcal{G}_e\hat{u}_e\right) + \tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}l_0^{-1}\dot{\tilde{\mathcal{W}}}_{c_0}. \tag{60}$$

Combining with (48a) and (52a), the following deductions are presented:

$$\sum_{i=1}^{M}\left(\nabla\mathcal{J}_i^{*\mathrm{T}}\left(\bar{\mathcal{F}}_i + \mathcal{G}_i\hat{u}_{p_i}\right)\right) = \sum_{i=1}^{M}\left(-\bar{z}_i^{\mathrm{T}}Q_i\bar{z}_i - \frac{1}{4}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i} + \frac{1}{2}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i} + \sigma_{c_i} + \varpi_i\right), \tag{61}$$

where $\varpi_i = \nabla\iota_{c_i}^{\mathrm{T}}(\bar{\mathcal{F}}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i})$. Similarly, using (48b) and (52b), the following equation is obtained:

$$-\nabla\mathcal{J}_0^{*\mathrm{T}}\sum_{i=1}^{M}\left(\mathcal{F}_i + \mathcal{G}_i\hat{u}_{p_i} + \mathcal{G}_e\hat{u}_e\right) = -z_0^{\mathrm{T}}Q_0z_0 - \frac{M^2}{4}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0} + \frac{M^2}{2}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0}$$
$$- \sum_{i=1}^{M}\frac{1}{2}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i} + \sigma_{c_0} + \varpi_0, \tag{62}$$

where $\varpi_0 = -\nabla\iota_{c_0}^{\mathrm{T}}\sum_{i=1}^{M}\left(\mathcal{F}_i - \frac{1}{2}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\hat{\mathcal{W}}_{c_i} + \frac{M}{2}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right)$.

In addition, the application of tuning laws (58a) and (58b) yields the following:

$$\sum_{i=1}^{M}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}l_i^{-1}\dot{\tilde{\mathcal{W}}}_{c_i} + \tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}l_0^{-1}\dot{\tilde{\mathcal{W}}}_{c_0}$$
$$= \sum_{i=1}^{M}\left\{\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\bar{\kappa}_i\left(-\bar{\kappa}_i^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i} + \frac{\sigma_{c_i}}{\hbar_i}\right) + \frac{1}{2}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\tilde{\mathcal{W}}_{c_i} - \frac{1}{4}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\mathcal{W}_{c_i}\right.$$
$$\left. - \frac{1}{2}\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\mathcal{W}_{c_0} + \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\left(\Upsilon_{i1}\hat{\mathcal{W}}_{c_i} - \Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\right) + \sigma_{c_i} + \varpi_i\right\} + \sigma_{c_0} + \varpi_0$$
$$+ \tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\bar{\kappa}_0\left(\bar{\kappa}_0^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0} + \frac{\sigma_{c_0}}{\hbar_0}\right) - \frac{M^2}{4}\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\mathcal{W}_{c_0} + \tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\left(\Upsilon_{01}\hat{\mathcal{W}}_{c_0} - \Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right). \tag{63}$$

Based on the definition of estimation errors in critic weights, given by $\tilde{\mathcal{W}}_{c_i} = \mathcal{W}_{c_i} - \hat{\mathcal{W}}_{c_i}$ and $\tilde{\mathcal{W}}_{c_0} = \mathcal{W}_{c_0} - \hat{\mathcal{W}}_{c_0}$, the following is derived:

$$\sum_{i=1}^{M}\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\left(\Upsilon_{i1}\hat{\mathcal{W}}_{c_i} - \Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\hat{\mathcal{W}}_{c_i}\right) = \sum_{i=1}^{M}\left(\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\Upsilon_{i1}\mathcal{W}_{c_i} - \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\Upsilon_{i1}\tilde{\mathcal{W}}_{c_i} - \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\mathcal{W}_{c_i} + \tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\Upsilon_{i2}\bar{\kappa}_i^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}\right), \tag{64}$$

$$\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\left(\Upsilon_{01}\hat{\mathcal{W}}_{c_0}-\Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\hat{\mathcal{W}}_{c_0}\right)=\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\Upsilon_{01}\mathcal{W}_{c_0}-\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\Upsilon_{01}\tilde{\mathcal{W}}_{c_0}-\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\mathcal{W}_{c_0}+\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\Upsilon_{02}\bar{\kappa}_0^{\mathrm{T}}\tilde{\mathcal{W}}_{c_0}. \tag{65}$$

For simplicity, denote $\mathcal{Y}_i^{\mathrm{T}}=[\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}}\bar{\kappa}_i,\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}\bar{\kappa}_0,\tilde{\mathcal{W}}_{c_i}^{\mathrm{T}},\tilde{\mathcal{W}}_{c_0}^{\mathrm{T}}]$ and combine with (61)–(65), which yields

$$\begin{aligned}\dot{V}_{\mathrm{HJI}}=&\sum_{i=1}^{M}\left\{-\mathcal{Y}_i^{\mathrm{T}}A_i\mathcal{Y}_i+\mathcal{Y}_i^{\mathrm{T}}B_i-\bar{z}_i^{\mathrm{T}}Q_i\bar{z}_i-\frac{1}{4}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}+\sigma_{c_i}+\varpi_i\right\}\\&-z_0^{\mathrm{T}}Q_0z_0-\frac{M^2}{4}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}+\sigma_{c_0}+\varpi_0,\end{aligned} \tag{66}$$

where

$$A_i=\begin{bmatrix}a_{i_{11}}&a_{i_{12}}&a_{i_{13}}&a_{i_{14}}\\a_{i_{21}}&a_{i_{22}}&a_{i_{23}}&a_{i_{24}}\\a_{i_{31}}&a_{i_{32}}&a_{i_{33}}&a_{i_{34}}\\a_{i_{41}}&a_{i_{42}}&a_{i_{43}}&a_{i_{44}}\end{bmatrix},\quad B_i=\begin{bmatrix}b_{i_1}\\b_{i_2}\\b_{i_3}\\b_{i_4}\end{bmatrix}$$

with the specific forms of the components in matrix A_i and vector B_i provided by the following:

$$\begin{aligned}&a_{i_{11}}=I,\ a_{i_{22}}=-I,\ a_{i_{33}}=\Upsilon_{i1},\ a_{i_{44}}=\Upsilon_{01},\ a_{i_{21}}=a_{i_{12}}^{\mathrm{T}}=a_{i_{41}}=a_{i_{14}}^{\mathrm{T}}=a_{i_{43}}=a_{i_{34}}^{\mathrm{T}}=0,\\&a_{i_{31}}=a_{i_{13}}^{\mathrm{T}}=-\frac{1}{2}\Upsilon_{i2},\ a_{i_{32}}=a_{i_{23}}^{\mathrm{T}}=\frac{1}{4\hbar_0}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0},\\&a_{i_{42}}=a_{i_{24}}^{\mathrm{T}}=-\frac{M^2}{4\hbar_0}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}-\frac{1}{2}\Upsilon_{02},\\&b_{i_1}=\frac{\sigma_{c_i}}{\hbar_i}-\Upsilon_{i2}\mathcal{W}_{c_i},\ b_{i_2}=\frac{\sigma_{c_0}}{\hbar_0}-\Upsilon_{02}\mathcal{W}_{c_0},\\&b_{i_3}=-\frac{1}{4}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}\frac{\bar{\kappa}_i^{\mathrm{T}}}{\hbar_i}\mathcal{W}_{c_i}+\Upsilon_{i1}\mathcal{W}_{c_i}+\frac{1}{2}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}-\frac{1}{2}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0},\\&b_{i_4}=-\frac{M}{4}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}\frac{\bar{\kappa}_0^{\mathrm{T}}}{\hbar_0}\mathcal{W}_{c_0}+\frac{M}{2}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}+\Upsilon_{01}\mathcal{W}_{c_0}.\end{aligned}$$

Considering Assumptions 5–7, which focus on the boundedness of the parameters and functions, the following conclusion is presented:

$$\begin{aligned}&\sum_{i=1}^{M}\left(-\frac{1}{4}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}+\sigma_{c_i}+\varpi_i\right)-\frac{M^2}{4}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}+\sigma_{c_0}+\varpi_0\\&=\sum_{i=1}^{M}\Big\{-\frac{1}{4}\mathcal{W}_{c_i}^{\mathrm{T}}\nabla\phi_{c_i}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\mathcal{W}_{c_i}+\frac{1}{2}\nabla\iota_{c_i}^{\mathrm{T}}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}+\frac{1}{4}\nabla\iota_{c_i}^{\mathrm{T}}\mathcal{D}_i\nabla\iota_{c_i}+\frac{1}{2}\nabla\iota_{c_0}^{\mathrm{T}}\mathcal{D}_i\nabla\phi_{c_i}^{\mathrm{T}}\tilde{\mathcal{W}}_{c_i}-\frac{1}{2}\nabla\iota_{c_0}^{\mathrm{T}}\mathcal{D}_i\nabla\iota_{c_i}\\&\quad-\frac{1}{2}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_i\nabla\iota_{c_i}\Big\}-\frac{M^2}{4}\mathcal{W}_{c_0}^{\mathrm{T}}\nabla\phi_{c_0}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}+\frac{M^2}{4}\nabla\iota_{c_0}^{\mathrm{T}}\mathcal{D}_0\nabla\iota_{c_0}+\frac{M^2}{2}\nabla\iota_{c_0}^{\mathrm{T}}\mathcal{D}_0\nabla\phi_{c_0}^{\mathrm{T}}\mathcal{W}_{c_0}\\&\leqslant\sum_{i=1}^{M}\left\{\left(\frac{1}{4}\bar{\phi}_{dc_i}^2\bar{\mathcal{W}}_{c_i}^2+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\phi}_{dc_i}\|\tilde{\mathcal{W}}_{c_i}\|+\frac{1}{4}\bar{\iota}_{dc_i}^2+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\iota}_{dc_0}+\frac{1}{2}\bar{\iota}_{dc_0}\bar{\phi}_{dc_i}\|\tilde{\mathcal{W}}_{c_i}\|+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\phi}_{dc_0}\bar{\mathcal{W}}_{c_0}\right)\lambda_{\max}(\mathcal{D}_i)\right\}\\&\quad+\frac{M^2}{4}\left(\bar{\phi}_{dc_0}^2\bar{\mathcal{W}}_{c_0}^2+\bar{\iota}_{dc_0}^2+2\bar{\iota}_{dc_0}\bar{\phi}_{dc_0}\bar{\mathcal{W}}_{c_0}\right)\lambda_{\max}(\mathcal{D}_0).\end{aligned} \tag{67}$$

By substituting (67) into (66), then

$$\begin{aligned}\dot{V}_{\mathrm{HJI}}&\leqslant\sum_{i=1}^{M}\left(-\lambda_{\min}(A_i)\|\mathcal{Y}_i\|^2+\|B_i\|\|\mathcal{Y}_i\|+\varrho_{m_i}-\lambda_{\min}(Q_i)\|\bar{z}_i\|^2\right)+\varrho_{m_0}-\lambda_{\min}(Q_0)\|z_0\|^2\\&\leqslant\sum_{i=1}^{M}\left\{-\lambda_{\min}(A_i)\left(\|\mathcal{Y}_i\|-\frac{\|B_i\|}{2\lambda_{\min}(A_i)}\right)^2-\lambda_{\min}(Q_i)\|\bar{z}_i\|^2\right\}-\lambda_{\min}(Q_0)\|z_0\|^2+\varrho_m,\end{aligned} \tag{68}$$

where

$$\varrho_{m_i}=\left(\frac{1}{4}\bar{\phi}_{dc_i}^2\bar{\mathcal{W}}_{c_i}^2+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\phi}_{dc_i}\|\tilde{\mathcal{W}}_{c_i}\|+\frac{1}{4}\bar{\iota}_{dc_i}^2+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\iota}_{dc_0}+\frac{1}{2}\bar{\iota}_{dc_0}\bar{\phi}_{dc_i}\|\tilde{\mathcal{W}}_{c_i}\|+\frac{1}{2}\bar{\iota}_{dc_i}\bar{\phi}_{dc_0}\bar{\mathcal{W}}_{c_0}\right)\lambda_{\max}(\mathcal{D}_i),$$

$$\varrho_{m_0} = \frac{M^2}{4}\left(\bar{\phi}_{dc_0}^2\bar{\mathcal{W}}_{c_0}^2 + \bar{\iota}_{dc_0}^2 + 2\bar{\iota}_{dc_0}\bar{\phi}_{dc_0}\bar{\mathcal{W}}_{c_0}\right)\lambda_{\max}(\mathcal{D}_0),$$
$$\varrho_m = \sum_{i=1}^{M}\left(\frac{\|B_i\|^2}{4\lambda_{\min}^2(A_i)} + \varrho_{m_i}\right) + \varrho_{m_0}.$$

Aiming to realize $\dot{V}_{\mathrm{HJI}} < 0$, one of the following relationships should hold:

$$\|\mathcal{Y}_i\| \geqslant \sqrt{\frac{\varrho_m}{\lambda_{\min}(A_i)}} + \frac{\|B_i\|}{2\lambda_{\min}(A_i)} \triangleq \Xi_i, \tag{69a}$$

$$\|\bar{z}_i\| \geqslant \sqrt{\frac{\varrho_m}{\lambda_{\min}(Q_i)}} \triangleq \Xi_{\bar{z}_i}, \tag{69b}$$

$$\|z_0\| \geqslant \sqrt{\frac{\varrho_m}{\lambda_{\min}(Q_0)}} \triangleq \Xi_{z_0}. \tag{69c}$$

By synthesizing this relationship with the Lyapunov stability theorem, $|\mathcal{Y}i|$ is uniformly ultimately bounded by Ξ_i, indicating that the estimated weight errors $\tilde{\mathcal{W}}c_i$ and $\tilde{\mathcal{W}}c_0$ are bounded. Consequently, $|\bar{z}_i|$ and $|z_0|$ are also uniformly ultimately bounded by Ξz_i and Ξ_{z_0}, respectively. Thus, the proof is complete.

Remark 5. Theorem 3 establishes uniform ultimate boundedness of the critic NN weights rather than exact convergence, ensuring bounded approximation errors and closed-loop stability. However, Assumptions 5–7 may be violated if state trajectories leave the compact approximation set, which could compromise stability. To address this, techniques such as state constraints, projection or saturation-based weight updates, and adaptive gain adjustment can be applied to enhance robustness.

4 Simulation studies

The simulations are conducted using the unicycle kinematic model, which is a standard representation for nonholonomic mobile agents such as ground vehicles or robots. The dynamics of each agent under (1) and (2) are formulated as

$$f(x_{p_i}) = 0_{3\times 1},\ g(x_{p_i}) = \begin{bmatrix}\cos(x_{p_{i3}}) & 0\\ \sin(x_{p_{i3}}) & 0\\ 0 & 1\end{bmatrix}, \tag{70}$$

$$f(x_e) = 0_{3\times 1},\quad g(x_e) = \begin{bmatrix}\cos(x_{e_3}) & 0\\ \sin(x_{e_3}) & 0\\ 0 & 1\end{bmatrix}, \tag{71}$$

where $x_{p_i} = [x_{p_{i1}}, x_{p_{i2}}, x_{p_{i3}}]^{\mathrm{T}}$ and $u_{p_i} = [u_{p_{i1}}, u_{p_{i2}}]^{\mathrm{T}}$ denote the state and control input of the i-th pursuer, while $x_e = [x_{e_1}, x_{e_2}, x_{e_3}]^{\mathrm{T}}$ and $u_e = [u_{e_1}, u_{e_2}]^{\mathrm{T}}$ correspond to those of the evader.

4.1 Case I: three-pursuer one-evader scenario

Consider an MPE game involving three pursuers and one evader, each governed by the unicycle dynamics described above. The communication links among agents follow a directed interaction topology, as illustrated in Figure 2.

Specifically, only Pursuer 1 directly receives information from the Evader. Pursuer 2 obtains information from Pursuer 1 with a nominal weight $\alpha_{21} = 1$, and Pursuer 3 obtains information from Pursuer 2. Communication link faults occur between the Evader and Pursuer 1, and between Pursuers 2 and 3, with time-varying weights modeled as $\bar{\gamma}_1(t) = 1 + 0.5\sin(t)$ and $\bar{\alpha}_{32}(t) = 1 + 0.3\sin(t)$.

Following the control design in Section 3, the simulation parameters are specified as follows. The observer parameters are chosen as $\hbar = 1$, $k = 0.5$, and $\beta_i(0) = 1$ for $i = 1, 2, 3$. To mitigate the chattering effect induced by the discontinuous sign function, $\mathrm{sgn}(\zeta_i)$ is replaced by its smooth approximation $\tanh(\zeta_i/\epsilon)$, with $\epsilon = 0.0005$. The weighting matrices in the performance index functions (24) and (25) are set to $Q_1 = Q_2 = Q_3 = 10I$, $Q_0 = -I$, and $R_1 = R_2 = R_3 = R_0 = 1$. The learning rates for the critic NNs are chosen as $l_1 = l_2 = l_3 = l_0 = 0.1$, and the activation functions are defined as follows $\phi_{c_i} = [z_{i_1}^2, z_{i_2}^2, 1-\cos(z_{i_3}), z_{i_1}z_{i_2}, z_{i_1}\sin(z_{i_3}), z_{i_2}\sin(z_{i_3})]^{\mathrm{T}}$, $i = 1, 2, 3$, $\phi_{c_0} = [z_{0_1}^2, z_{0_2}^2, 1-\cos(z_{0_3}), z_{0_1}z_{0_2}, z_{0_1}\sin(z_{0_3}), z_{0_2}\sin(z_{0_3})]^{\mathrm{T}}$. The estimated weights for the critic networks are denoted as follows $\hat{\mathcal{W}}_{c_i} = [\hat{\mathcal{W}}_{c_{i1}}, \hat{\mathcal{W}}_{c_{i2}}, \hat{\mathcal{W}}_{c_{i3}}, \hat{\mathcal{W}}_{c_{i4}}, \hat{\mathcal{W}}_{c_{i5}}, \hat{\mathcal{W}}_{c_{i6}}]^{\mathrm{T}}$, $i = 1, 2, 3$, $\hat{\mathcal{W}}_{c_0} = [\hat{\mathcal{W}}_{c_{01}}, \hat{\mathcal{W}}_{c_{02}}, \hat{\mathcal{W}}_{c_{03}}, \hat{\mathcal{W}}_{c_{04}}, \hat{\mathcal{W}}_{c_{05}}, \hat{\mathcal{W}}_{c_{06}}]^{\mathrm{T}}$. All

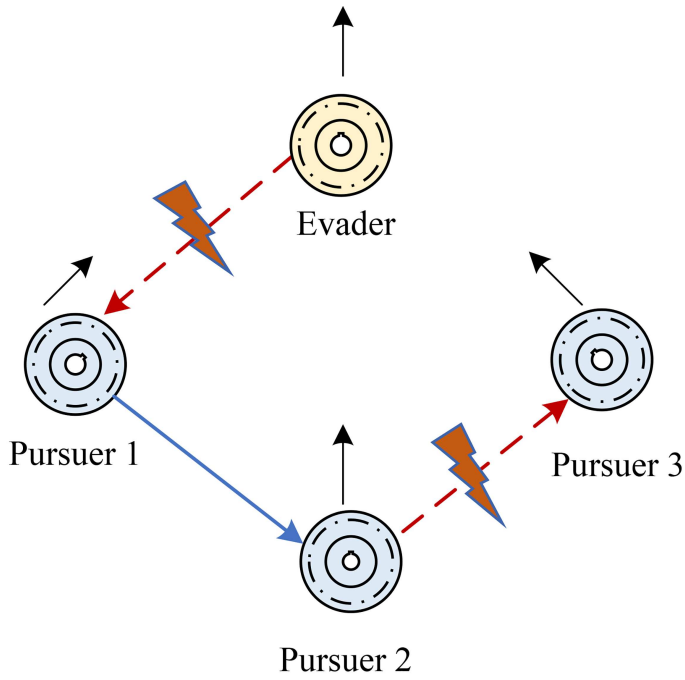

Figure 2 (Color online) MPE game with three pursuers and one evader under communication link faults.

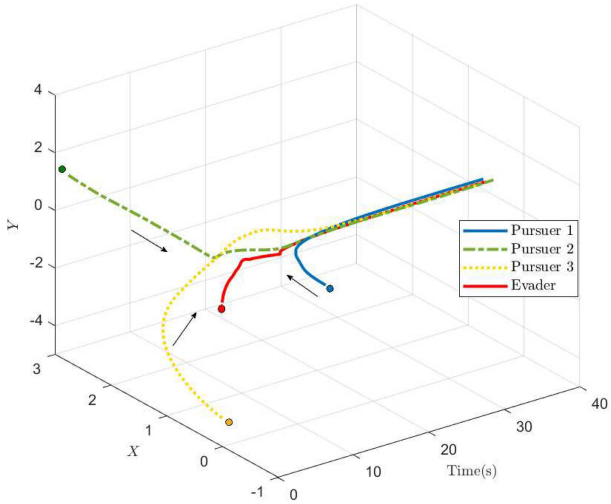

Figure 3 (Color online) Trajectories of the MPE game with three pursuers and one evader.

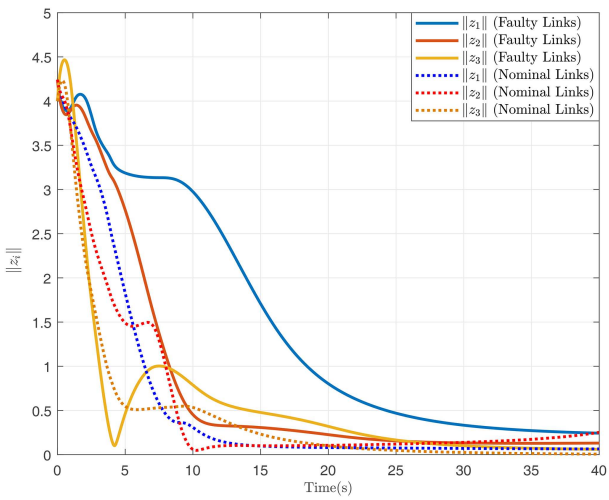

Figure 4 (Color online) State errors between three pursuers and one evader under faulty and nominal links.

initial weights $\hat{\mathcal{W}}_{c_i}$ and $\hat{\mathcal{W}}_{c_0}$ are randomly initialized within the range $[0, 1]$. Additionally, the initial state of the system is set as follows $x_{p_1}(0) = [-3, 3, \frac{\pi}{4}]$, $x_{p_2}(0) = [3, 3, \frac{3\pi}{4}]$, $x_{p_3}(0) = [0, -4, 0]$, and $x_e(0) = [0, 0, \frac{\pi}{2}]$.

The simulation results in Figures 3–10 verify the effectiveness of the proposed method in handling the MPE game under communication link faults. As depicted in Figure 3, the three pursuers successfully intercept the evader within approximately 20 s. Figure 4 shows that the relative state errors converge to zero almost simultaneously regardless of link faults, with larger oscillations and slower convergence under faulty links effectively mitigated by the observer's adaptive gain, keeping all errors bounded. Figures 5–8 illustrate the convergence of the critic network weights, which stabilize within the desired range. Figure 9 depicts the adaptive gain evolution in the distributed observer, reflecting its ability to adjust dynamically to varying system conditions. Finally, Figure 10 demonstrates the optimized strategies of the pursuers, exhibiting the capability of the critic network to approximate the Nash equilibrium and optimize pursuit strategies in real time. These results collectively highlight the robustness and adaptability of the proposed method.

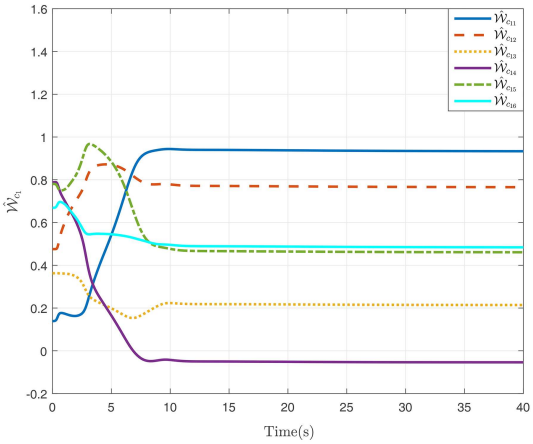

Figure 5 (Color online) Convergence curves of the estimated weight $\hat{\mathcal{W}}_{c_1}$.

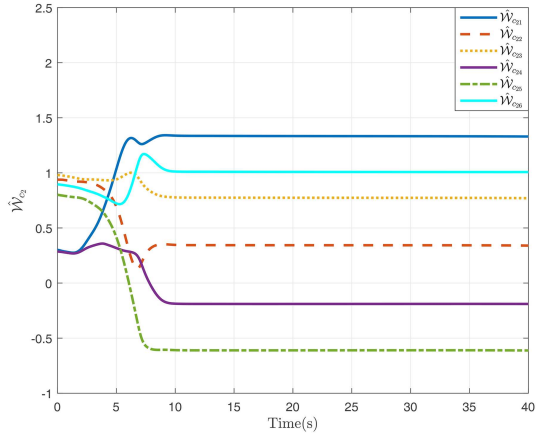

Figure 6 (Color online) Convergence curves of the estimated weight $\hat{\mathcal{W}}_{c_2}$.

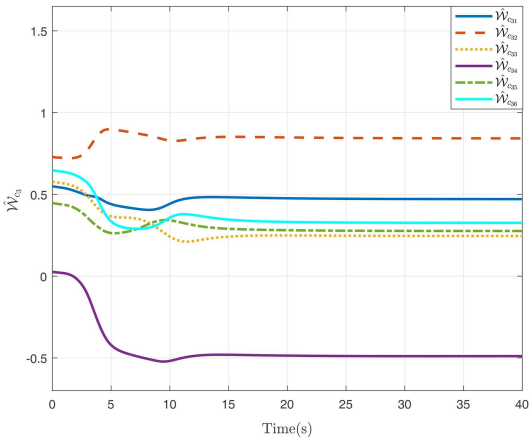

Figure 7 (Color online) Convergence curves of the estimated weight $\hat{\mathcal{W}}_{c_3}$.

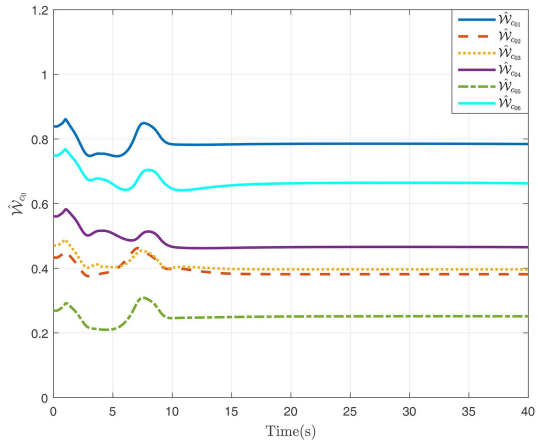

Figure 8 (Color online) Convergence curves of the estimated weight $\hat{\mathcal{W}}_{c_0}$.

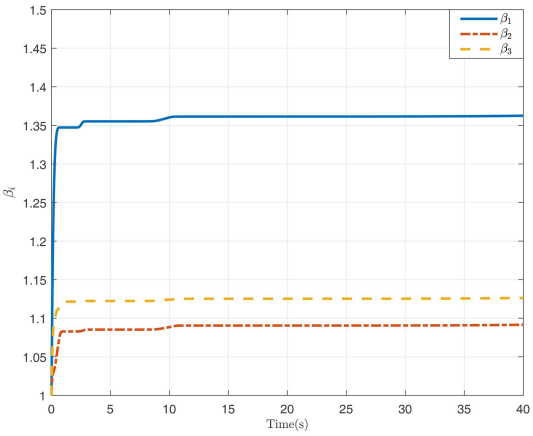

Figure 9 (Color online) Trajectories of adaptive observer gains β_1, β_2, β_3.

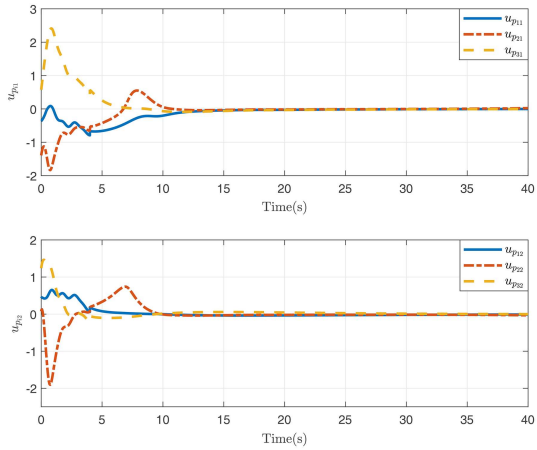

Figure 10 (Color online) Evolution of the optimal pursuit strategies $u_{p_{i1}}$, $u_{p_{i2}}$ for all pursuers ($i = 1, 2, 3$).

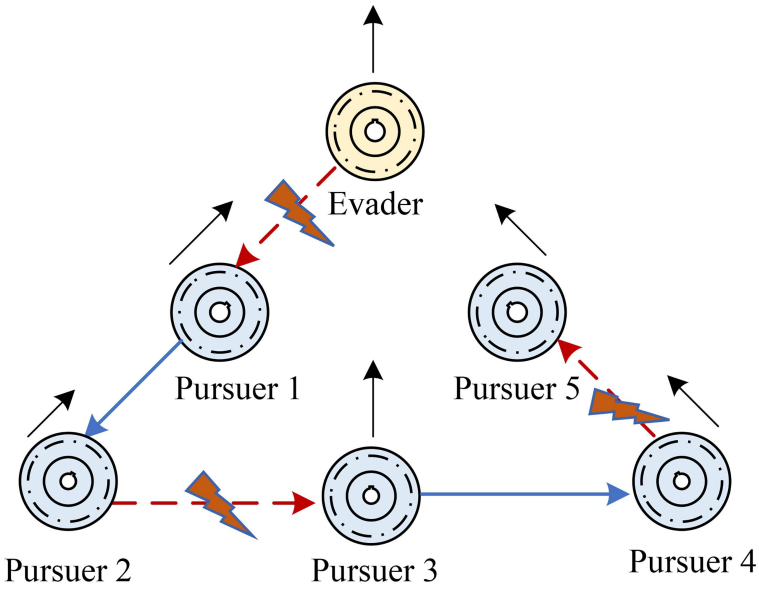

Figure 11 (Color online) MPE game with five pursuers and one evader under communication link faults.

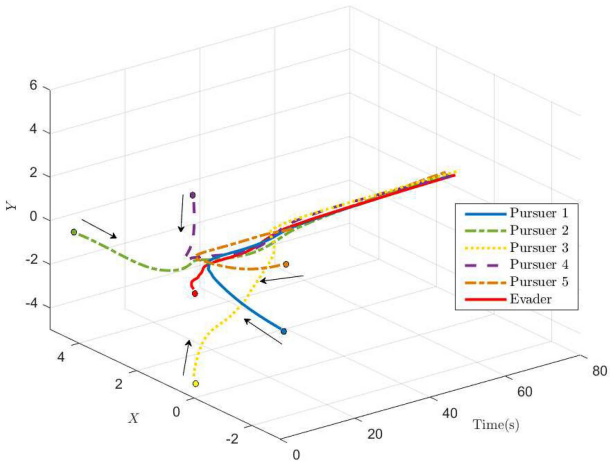

Figure 12 (Color online) Trajectories of the MPE game with five pursuers and one evader.

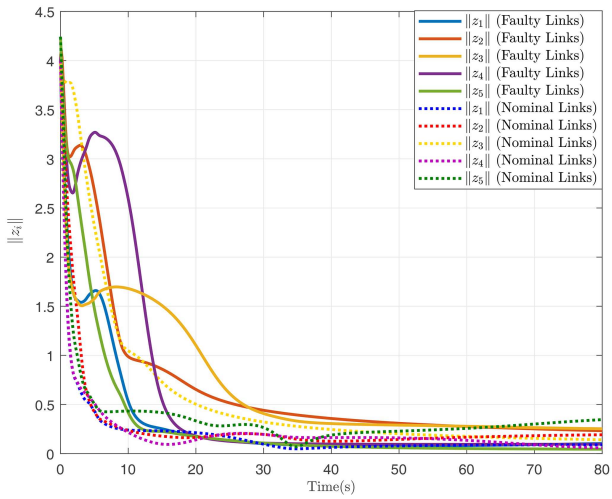

Figure 13 (Color online) State errors between five pursuers and one evader under faulty and nominal links.

4.2 Case II: five-pursuer one-evader scenario

To further examine the scalability of the proposed strategy, a second scenario considers five pursuers and one evader. The directed communication topology among agents is illustrated in Figure 11.

The communication topology is configured such that Pursuer 1 directly accesses the Evader, Pursuer 2 receives information from Pursuer 1 with $\alpha_{21} = 1$, Pursuer 3 from Pursuer 2, Pursuer 4 from Pursuer 3 with $\alpha_{43} = 1$, and Pursuer 5 from Pursuer 4. Link faults occur between the Evader and Pursuer 1, between Pursuers 2 and 3, and between Pursuers 4 and 5, with time-varying weights $\bar{\gamma}_1(t) = 1 + 0.5\sin(t)$, $\bar{\alpha}_{32}(t) = 1 + 0.3\sin(t)$, and $\bar{\alpha}_{54}(t) = 1 + 0.3\sin(t)$.

To ensure a fair comparison between the two cases, the control and observer settings are consistent with Section 3. Specifically, the observer parameters are chosen as $\hbar = 2$, $k = 1$, and $\beta_i(0) = 1$. To mitigate chattering, the discontinuous $\mathrm{sgn}(\zeta_i)$ is replaced by $\tanh(\zeta_i/\epsilon)$ with $\epsilon = 0.0001$. The weighting matrices in (24) and (25) are selected as $Q_1 = Q_2 = Q_3 = Q_4 = Q_5 = 10I$, $Q_0 = -I$, and $R_1 = R_2 = R_3 = R_4 = R_5 = R_0 = 1$. The critic networks use a learning rate of 0.1, identical activation functions as in Case I, and randomly initialized weights in $[0, 1]$. The pursuers start from $x_{p_1}(0) = [-4, 0, 0]$, $x_{p_2}(0) = [4, 0, \pi]$, $x_{p_3}(0) = [0, -4, \frac{\pi}{2}]$, $x_{p_4}(0) = [0, 4, -\frac{\pi}{2}]$, and $x_{p_5}(0) = [-3, 3, -\frac{\pi}{4}]$, while the evader begins at $x_e(0) = [0, 0, \frac{\pi}{4}]$. This configuration tests the scalability of the

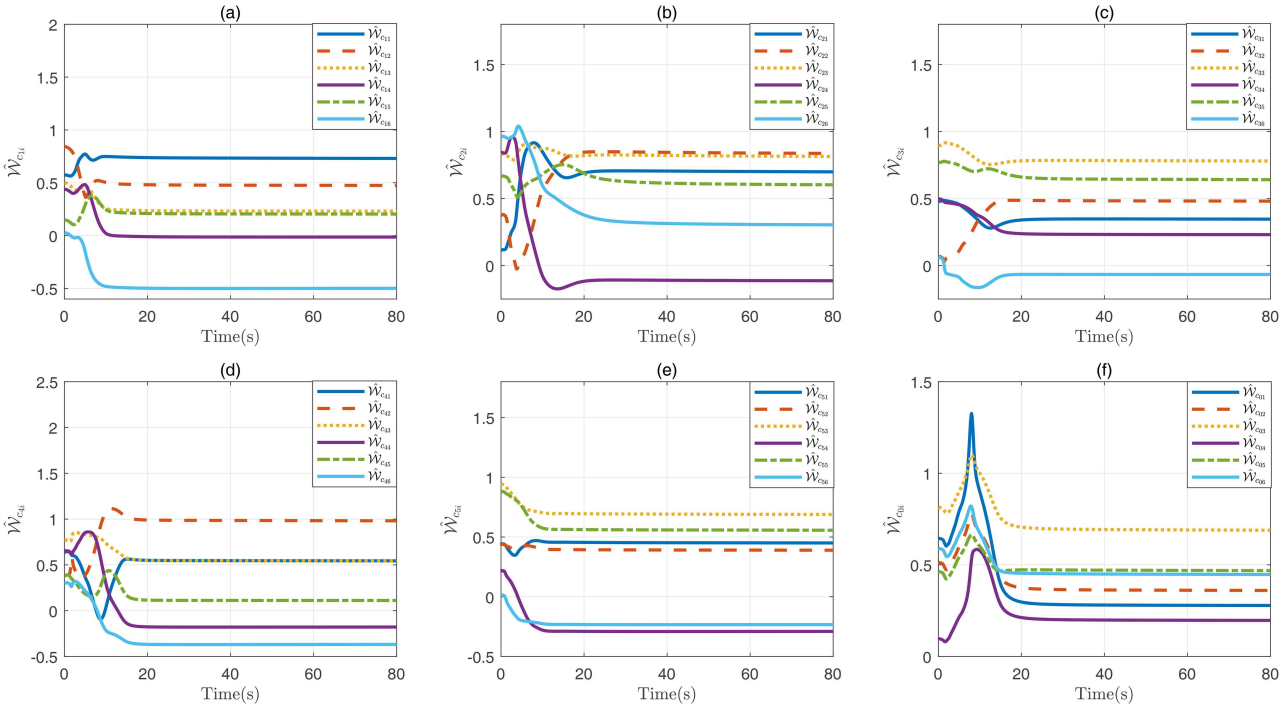

Figure 14 (Color online) Convergence curves of the estimated weights (a) $\hat{\mathcal{W}}_{c_1}$, (b) $\hat{\mathcal{W}}_{c_2}$, (c) $\hat{\mathcal{W}}_{c_3}$, (d) $\hat{\mathcal{W}}_{c_4}$, (e) $\hat{\mathcal{W}}_{c_5}$, (f) $\hat{\mathcal{W}}_{c_0}$.

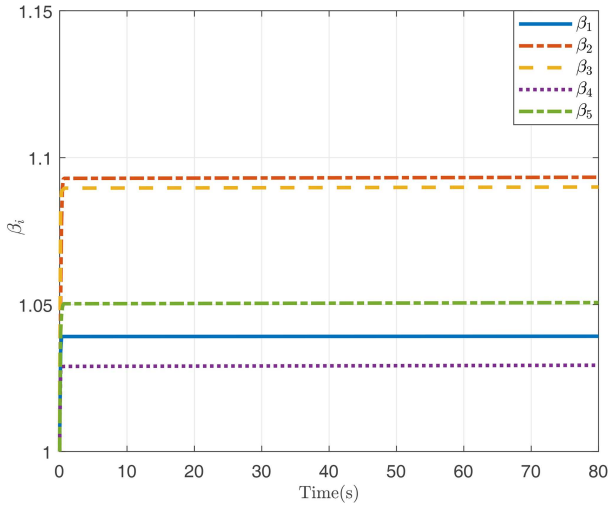

Figure 15 (Color online) Trajectories of adaptive observer gains β_1, β_2, β_3, β_4, β_5.

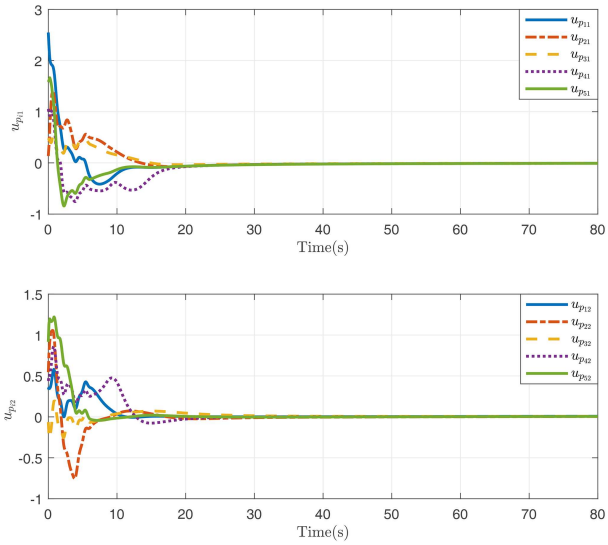

Figure 16 (Color online) Evolution of the optimal pursuit strategy u_{pi1}, u_{pi2} for all pursuers ($i = 1, 2, 3, 4, 5$).

proposed method as the number of pursuers increases beyond three.

Building upon the preceding scenario, we consider a more challenging case with five pursuers and one evader, initialized at different positions. Communication link failures are assumed between Pursuers 2 and 3, and 4 and 5. Figures 12–16 show the agent trajectories, capture errors, and control inputs. Despite these link faults, the distributed fault-tolerant strategy enables collaborative tracking and capture of the evader within a finite time, while maintaining overall system stability. For clarity of analysis, the simulations are conducted in an open environment to isolate the effects of communication link faults; future work will consider obstacles or disturbances. This case also demonstrates the scalability of the proposed method: each pursuer relies only on local neighbor information, keeping per-agent computation low, while communication load grows with network size. By leveraging sparse or structured topologies, the approach can handle larger groups while preserving robustness against both actuator and

communication faults, confirming its effectiveness in more complex MPE scenarios.

5 Conclusion

In this paper, a novel FTC framework has been proposed for MPE games under communication link faults. A distributed adaptive observer was designed to estimate the evader's states despite time-varying and uncertain topologies, while a single-critic NN enabled efficient real-time approximation of optimal strategies by solving the HJI equations. Simulations on the nonlinear unicycle model verified the practicality and effectiveness of the method. Future work will extend the framework to more realistic scenarios involving intermittent link failures, delays, adversarial disturbances, and environments with obstacles to further enhance robustness and applicability.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62225303, 62433004) and Fundamental Research Funds for the Central Universities (Grant Nos. buctrc202228, buctrc202201).

References

1 Zhang R, Zong Q, Zhang X, et al. Game of drones: multi-UAV pursuit-evasion game with online motion planning by deep reinforcement learning. IEEE Trans Neural Netw Learn Syst, 2023, 34: 7900–7909
2 Li F, Yin M, Wang T, et al. Distributed pursuit-evasion game of limited perception USV swarm based on multiagent proximal policy optimization. IEEE Trans Syst Man Cybern Syst, 2024, 54: 6435–6446
3 Zhang Z, Zhang K, Xie X, et al. Fixed-time zero-sum pursuit-evasion game control of multisatellite via adaptive dynamic programming. IEEE Trans Aerosp Electron Syst, 2024, 60: 2224–2235
4 Zheng Z, Zhang P, Yuan J. Nonzero-sum pursuit-evasion game control for spacecraft systems: a Q-learning method. IEEE Trans Aerosp Electron Syst, 2023, 59: 3971–3981
5 Ho Y, Bryson A, Baron S. Differential games and optimal pursuit-evasion strategies. IEEE Trans Automat Contr, 1965, 10: 385–389
6 Meier L. A new technique for solving pursuit-evasion differential games. IEEE Trans Automat Contr, 1969, 14: 352–359
7 Oyler D W, Kabamba P T, Girard A R. Pursuit-evasion games in the presence of obstacles. Automatica, 2016, 65: 1–11
8 Li W. A dynamics perspective of pursuit-evasion: capturing and escaping when the pursuer runs faster than the agile evader. IEEE Trans Automat Contr, 2017, 62: 451–457
9 Bu S M, Liang L, Wang Y Q. A mixed Nash equilibrium solution for visibility-based pursuit-evasion game with multiple obstacles. Sci China Inf Sci, 2022, 67: 194201
10 Foley M, Schmitendorf W. A class of differential games with two pursuers versus one evader. IEEE Trans Automat Contr, 1974, 19: 239–243
11 Zhang F, Zha W Z. Evasion strategies of a three-player lifeline game. Sci China Inf Sci, 2018, 61: 112206
12 Chen J, Zha W, Peng Z, et al. Multi-player pursuit-evasion games with one superior evader. Automatica, 2016, 71: 24–32
13 Von Moll A, Pachter M, Garcia E, et al. Robust policies for a multiple-pursuer single-evader differential game. Dyn Games Appl, 2020, 10: 202–221
14 Selvakumar J, Bakolas E. Feedback strategies for a reach-avoid game with a single evader and multiple pursuers. IEEE Trans Cybern, 2021, 51: 696–707
15 Fang X, Wang C, Xie L, et al. Cooperative pursuit with multi-pursuer and one faster free-moving evader. IEEE Trans Cybern, 2022, 52: 1405–1414
16 Wang L B, Liu Z X, Yuan S, et al. Distributed Nash equilibrium for pursuit-evasion game with one evader and multiple pursuers. Sci China Inf Sci, 2025, 68: 192205
17 Chen C, Xie K, Lewis F L, et al. Adaptive synchronization of multi-agent systems with resilience to communication link faults. Automatica, 2020, 111: 108636
18 Zhao W, Liu H, Valavanis K P, et al. Fault-tolerant formation control for heterogeneous vehicles via reinforcement learning. IEEE Trans Aerosp Electron Syst, 2022, 58: 2796–2806
19 Gong J, Jiang B, Ma Y, et al. Distributed adaptive fault-tolerant formation control for heterogeneous multiagent systems with communication link faults. IEEE Trans Aerospace Electron Syst, 2023, 59: 784–795
20 Liu G, Sun Q, Wang R, et al. Reduced-order observer-based fuzzy adaptive dynamic event-triggered consensus control for multi-agent systems with communication faults. Nonlinear Dyn, 2022, 110: 1421–1435
21 Sun J, Xu Z, Zhang H, et al. Adaptive distributed control of nonlinear multiagent systems with event-triggered for communication faults and dead-zone inputs. IEEE Trans Cybern, 2024, 54: 5877–5886
22 Su H S, Li Z H. Adaptive leader-following attitude consensus of multiple rigid body systems with resilience to communication link faults. Sci China Inf Sci, 2023, 66: 209201
23 Gong J, Ma Y, Jiang B, et al. An adaptive fault-tolerant control scheme for heterogeneous multiagent systems. IEEE Trans Syst Man Cybern Syst, 2025, 55: 1264–1276
24 He C, Qi R, Jiang B. Fixed-time adaptive neural tracking control for nonlinear multiagent systems with communication link faults. Intl J Robust Nonlinear, 2024, 34: 1799–1827
25 Jagat A, Sinclair A J. Nonlinear control for spacecraft pursuit-evasion game using the state-dependent Riccati equation method. IEEE Trans Aerosp Electron Syst, 2017, 53: 3032–3042
26 Xiong H, Zhang Y. Reinforcement learning-based formation-surrounding control for multiple quadrotor UAVs pursuit-evasion games. ISA Trans, 2024, 145: 205–224
27 Vamvoudakis K G, Lewis F L, Hudas G R. Multi-agent differential graphical games: online adaptive learning solution for synchronization with optimality. Automatica, 2012, 48: 1598–1611
28 Song R, Wei Q, Song B. Neural-network-based synchronous iteration learning method for multi-player zero-sum games. Neurocomputing, 2017, 242: 73–82
29 Jiang H, Zhang H, Han J, et al. Iterative adaptive dynamic programming methods with neural network implementation for multi-player zero-sum games. Neurocomputing, 2018, 307: 54–60
30 Gong Z, He B, Hu C, et al. Online adaptive dynamic programming-based solution of networked multiple-pursuer and single-evader game. Electronics, 2022, 11: 3583
31 Liu M, Liu Q H, Zhang L X, et al. Adaptive dynamic programming-based fault-tolerant attitude control for flexible spacecraft with limited wireless resources. Sci China Inf Sci, 2023, 66: 202201
32 Wang Y, Hou W, Liang L. Prescribed performance optimal fault-tolerant control for nonlinear systems with mismatched disturbances via zero-sum differential game. ISA Trans, 2025, 164: 91–99
33 Meng Y, Liu C, Wang Q, et al. Cooperative advantage actor-critic reinforcement learning for multiagent pursuit-evasion games on communication graphs. IEEE Trans Artif Intell, 2024, 5: 6509–6523

34 Qiao J, Li M, Wang D. Asymmetric constrained optimal tracking control with critic learning of nonlinear multiplayer zero-sum games. IEEE Trans Neural Netw Learn Syst, 2024, 35: 5671–5683
35 Xu Z, Yu D, Liu Y J, et al. Approximate optimal strategy for multiagent system pursuit-evasion game. IEEE Syst J, 2024, 18: 1669–1680
36 Yuan Y, Zhang P, Li X. Synchronous fault-tolerant near-optimal control for discrete-time nonlinear PE game. IEEE Trans Neural Netw Learn Syst, 2021, 32: 4432–4444
37 Lopez V G, Lewis F L, Wan Y, et al. Solutions for multiagent pursuit-evasion games on communication graphs: finite-time capture and asymptotic behaviors. IEEE Trans Automat Contr, 2020, 65: 1911–1923
38 Beke A, Kumbasar T. Type-2 fuzzy logic-based linguistic pursuing strategy design and its deployment to a real-world pursuit evasion game. IEEE Trans Cybern, 2020, 50: 211–221
39 Xu Y, Yang H, Jiang B, et al. Multiplayer pursuit-evasion differential games with malicious pursuers. IEEE Trans Automat Contr, 2022, 67: 4939–4946
40 Cheng L, Yuan Y. Multiplayer obstacle avoidance pursuit-evasion games with adaptive parameter estimation. IEEE Trans Ind Electron, 2023, 70: 5171–5181
41 Zhou P, Chen B M. Distributed optimal solutions for multiagent pursuit-evasion games for capture and formation control. IEEE Trans Ind Electron, 2024, 71: 5224–5234
42 Wang Y, Zhang H, Wang Z, et al. Nash-minmax strategies for multiagent pursuit-evasion games with reinforcement learning. IEEE Trans Cybern, 2025, 55: 3320–3331