Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Safety-gated graph-attention compensation backstepping control for multi-UAV formation flight in obstacle-rich environments

Hao Wang
Air Force Engineering University
Guhao Zhao
zhaoguhao_1986@163.com

Air Force Engineering University
Yulong Cao
Air Force Engineering University

Article

Keywords: Multi-UAV formation control, Graph attention mechanism, Safety gating, Backstepping control, Obstacle avoidance

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11588136/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Safety-gated graph-attention compensation backstepping control for multi-UAV formation flight in obstacle-rich environments

Hao Wang[1], Guhao Zhao[1*], Yulong Cao[1]

[1*]Air Traffic Control and Navigation School, Air Force Engineering University, Xi'an, 710051, Shaanxi, China.

*Corresponding author(s). E-mail(s): zhaoguhao_1986@163.com;
Contributing authors: wh05072025@163.com; caoyulong_1998@163.com;

Abstract

This paper develops a safety-gated graph-attention compensation backstepping architecture for multi-UAV formation flight among static and dynamic obstacles. Each follower is represented by a three-dimensional double-integrator model with bounded acceleration. A sliding-surface law tracks the desired formation; an obstacle-distance authority factor and a safety-conflict projection coordinate tracking and avoidance; a graph attention network (GAT) generates risk-aware additive compensation; and a safety gate redistributes authority among tracking, obstacle avoidance, inter-UAV anti-collision, altitude constraint, and compensation channels. Smooth saturation enforces the input bound. A Lyapunov analysis of the smooth nominal control law establishes uniform ultimate boundedness in the unsaturated operating region, while restart-enabled covariance matrix adaptation evolution strategy (CMA-ES) tunes the ten coupled gains offline. In a dense dynamic scene, the tuned controller attains 90.0% obstacle-avoidance success, 66.7% joint success, a mean minimum obstacle clearance of $\mathbf{0.249 \pm 0.020}$ m, and a tracking RMSE of $\mathbf{0.879}$ m. Removing GAT compensation increases RMSE to $\mathbf{2.40}$ m, and CMA-ES tuning increases dense-scene joint success from **3.3%** to **66.7%**. Ablations identify obstacle avoidance and predictive avoidance as safety-critical modules, nominal tracking and GAT compensation as performance-critical modules, and the residual inter-UAV margin as the dominant remaining bottleneck.

Keywords: Multi-UAV formation control, Graph attention mechanism, Safety gating, Backstepping control, Obstacle avoidance

1 Introduction

Multi-UAV formations are increasingly used in inspection, search-and-rescue, logistics, and surveillance [1–3]. In cluttered airspace, each vehicle must simultaneously track a time-varying formation, avoid static and dynamic obstacles, maintain separation from neighbors, and reject disturbances under bounded acceleration. These objectives are strongly coupled: an evasive maneuver enlarges formation error, whereas aggressive formation restoration can reduce obstacle or inter-UAV clearance. Safe formation flight is therefore a closed-loop control problem as well as a trajectory-generation problem.

Optimization-based controllers constitute the first family. Dentler et al. combined model predictive control with cooperative localization under potential-function sensor constraints [1]; Wang et al. and Jiang et al. developed distributed consensus and formation protocols for large-scale swarms [2, 3]; Dai et al. integrated real-time collision-free planning with tracking control [4]; Niu et al. designed a communication-free MPC to reduce dependence on the communication topology [5]; Huo et al. realized collision-free model-predictive trajectory tracking [6]; Niu et al. (2025) achieved agile avoidance through dynamic velocity-obstacle perception [7]; and Chen et al. coupled reference-path planning with MPC [8]. These schemes handle constraints explicitly through receding-horizon optimization, but they rely on model accuracy, incur nontrivial online computation, depend on manually tuned weights, and may respond late to fast-moving obstacles.

Artificial-potential-field methods form the second family. Zhu et al. proposed a flexible collision-avoidance strategy for UAV formations [9], while Zhang et al. and Ma et al. refined the potential functions for dense multi-UAV scenes [10, 11], and Li et al. addressed obstacle avoidance and formation control in unknown environments [12]. APF is structurally simple and real-time friendly, but it suffers from local minima and goal-unreachability, requires manual repulsion-gain tuning, and tends to induce command saturation and oscillation in dense dynamic fields; moreover, the authority conflict between repulsion and formation tracking lacks a systematic coordination mechanism. Geometric conflict-resolution methods [13] adjust heading and elevation angles from geometric relations and miss-distance criteria; they usually assume constant or simple motion and couple weakly with the low-level formation law, which limits their applicability under aggressive maneuvers.

Control barrier functions form the third family and enforce safety explicitly: Lv et al. derived CBF-based collision-avoidance guidance for multi-fixed-wing UAVs [14], Fu et al. combined robust obstacle avoidance with safe formation tracking [15], and Wei et al. addressed formation safety under switched trajectories and switching communication topologies [16]. However, the number of pairwise constraints grows quadratically with fleet size, multiple constraints may conflict in feasibility, and the online trade-off between safety and tracking performance is difficult to coordinate. Learning-based navigation and planning form the fourth family: Liu et al. proposed a hierarchical reinforcement-learning algorithm with an attention mechanism [17], and subsequent attention- or deep-model-based studies [18–20] improved multi-agent interaction modeling and cooperative decision-making. These methods capture complex interactions, but they output a policy or plan without a closed-loop low-level safety guarantee and face simulation-to-reality transfer difficulties. Finally, backstepping and sliding-mode

controllers provide mature convergence design tools: Zhang et al. realized multi-UAV formation control via backstepping, and Muslimov et al. realized adaptive flocking control via vector fields and backstepping [21, 22], Shi et al. addressed tracking and obstacle avoidance of a high-order fully actuated quadrotor [23], and Li et al. designed an adaptive backstepping sliding-mode tracker for quadrotors [24]. Yet these controllers offer limited explicit awareness of neighbor risk, treat avoidance through outer-loop planning or simple repulsion, and switch stiffly among tracking, avoidance, and anti-collision channels, so safety margins are difficult to guarantee in dense scenes.

Overall, three gaps emerge from the control and obstacle-avoidance perspective. First, perception and control are decoupled: attention mechanisms and learned models mostly remain at the decision level, so the low-level control law cannot exploit neighbor-risk information for online authority coordination. Second, multi-objective authority conflicts lack a systematic mechanism: channel weights among tracking, avoidance, and anti-collision are set by manual experience, switching is non-smooth, and commands tend to chatter or saturate in dense dynamic scenes. Third, coupled gain tuning is difficult: gains of multi-channel control laws are strongly coupled, and manually balancing safety against performance is impractical.

Against these gaps, this paper proposes a safety-gated graph-attention compensation backstepping controller for multi-UAV formation flight in obstacle-rich environments. The graph attention network (GAT) [25] acts as a bounded additive compensation term inside the backstepping law, while a safety gate redistributes control authority online among the tracking, obstacle-avoidance, inter-UAV anti-collision, altitude-constraint, and compensation channels according to obstacle and inter-UAV risk. The ten coupled gains are tuned offline with restart-enabled CMA-ES [26–28]; CMA-ES is a reproducible tuning tool rather than the central contribution.

The main contributions are threefold:

(a) **Control architecture.** An attention-as-compensation formation architecture with safety-gated authority redistribution: the complementary pair ($g_{\mathrm{gate},i}$, w_z) hands authority smoothly between the compensation and avoidance channels, keeping commands chattering-free and free of sustained saturation under input limits.

(b) **Obstacle-avoidance mechanism.** A layered safety-margin regulation strategy that combines 0.5 s predictive repulsion to remove dynamic-obstacle lag, an obstacle-distance authority factor that simultaneously reduces tracking gain and increases repulsion near obstacles, and smooth switching with distance truncation against command singularity.

(c) **Experimental validation.** A unified 30-seed, two-scenario protocol shows that pure backstepping violates obstacle safety, an APF baseline saturates into infeasibility, and removing GAT compensation degrades the tracking RMSE to 2.40 m, whereas CMA-ES tuning raises the dense-scene joint success rate from 3.3% to 66.7%; the ablation study further identifies the residual inter-UAV margin as the dominant bottleneck.

2 Modeling and Control Design of the Multi-UAV Formation System

2.1 UAV Dynamic Model and Formation Model

In this work, we study a UAV swarm that performs formation flight in a three-dimensional obstacle-filled environment. The swarm consists of one leader UAV and $N = 4$ follower UAVs. The leader follows a nominal pre-planned profile and an obstacle-aware leader-side correction that is independent of the followers. During the flight, each follower must satisfy three constraints:

1. Maintain the desired formation offset relative to the leader UAV;
2. Avoid static obstacles and dynamic obstacles during operation;
3. Prevent inter-UAV collisions during flight.

The position and velocity of the i-th UAV are

$$\boldsymbol{p}_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{v}_i = \begin{bmatrix} v_{x,i} & v_{y,i} & v_{z,i} \end{bmatrix}^{\mathrm{T}} \tag{1}$$

The control input is a three-dimensional acceleration command $\boldsymbol{u}_i = \begin{bmatrix} u_{x,i} & u_{y,i} & u_{z,i} \end{bmatrix}^{\mathrm{T}}$. Each UAV follows a three-dimensional double-integrator point-mass model

$$\begin{aligned} \dot{\boldsymbol{p}}_i &= \boldsymbol{v}_i, \\ \dot{\boldsymbol{v}}_i &= \boldsymbol{u}_i + \boldsymbol{\delta}_i(t), \end{aligned} \tag{2}$$

and its discretized implementation uses the semi-implicit Euler update with velocity and altitude projections,

$$\begin{aligned} \boldsymbol{v}_i[k+1] &= \Pi_v \left[\boldsymbol{v}_i[k] + \big(\boldsymbol{u}_i[k] + \boldsymbol{\delta}_i[k] \big) \Delta t \right], \\ \boldsymbol{p}_i[k+1] &= \Pi_z [\boldsymbol{p}_i[k] + \boldsymbol{v}_i[k+1] \Delta t], \end{aligned} \tag{3}$$

where Π_v first imposes the three-dimensional speed bound $\|\boldsymbol{v}_i\| \le v_{\max}$ and then clips $|v_{z,i}| \le v_{z,\max}$; Π_z clips altitude to the numerical containment interval $[z_{\min} - 0.25, z_{\max} + 0.25]$ m. The softer altitude-control corridor $[z_{\min}, z_{\max}]$ is enforced by the control channel before this defensive projection. Here $\boldsymbol{p}_i$ and $\boldsymbol{v}_i$ denote the position and velocity, respectively, $\boldsymbol{u}_i$ is the acceleration command to design, and Δt is the simulation step. The term $\boldsymbol{\delta}_i(t) \in \mathbb{R}^3$ denotes the lumped disturbance, with $\|\boldsymbol{\delta}_i(t)\| \le \delta_{\max}$ in the analysis. In the simulations it combines a heterogeneous time-varying component, intermittent impulses, and discrete acceleration white noise.

This paper adopts a leader-follower formation architecture with one leader and N followers. The actual leader state $\big(\boldsymbol{p}_L(t), \boldsymbol{v}_L(t), \boldsymbol{a}_L(t)\big)$ follows a nominal pre-planned profile with an obstacle-aware leader-side predictive correction; it is generated independently of the followers and broadcast to them. We define the desired formation position of the i-th follower as

$$\boldsymbol{p}_{d,i}(t) = \boldsymbol{p}_L(t) + \Delta \boldsymbol{d}_i \tag{4}$$

Here $\Delta \boldsymbol{d}_i \in \mathbb{R}^3$ denotes the predefined formation offset of the follower. In this work, the four followers form a rear-layered rhombus arranged in two rows: the first row lies closer to the leader and in the upper layer, and the second row lies farther behind and in the lower layer, with left–right symmetry within each row. The offsets are chosen so that the initial follower–follower separations lie inside the communication range R_{comm} and above $d_{\text{uav}}^{\text{safe}}$; their numerical values are specified in Section 3.1. The desired velocity of each follower equals the leader velocity:

$$\boldsymbol{v}_{d,i}(t) = \boldsymbol{v}_L(t) \tag{5}$$

According to the kinematic and formation models, the position and velocity errors of the i-th UAV take the form

$$\begin{aligned} \boldsymbol{e}_{p,i} &= \boldsymbol{p}_i - \boldsymbol{p}_{d,i}, \\ \boldsymbol{e}_{v,i} &= \boldsymbol{v}_i - \boldsymbol{v}_{d,i}. \end{aligned} \tag{6}$$

where $\boldsymbol{e}_{p,i} \in \mathbb{R}^3$ and $\boldsymbol{e}_{v,i} \in \mathbb{R}^3$ each contain the components along the three coordinate axes.

2.2 Graph Modeling of UAV Swarm

This paper introduces graph attention networks into the UAV swarm framework. They compute a risk-aware attention score for each UAV, which supports the cooperative compensation in the next section. Crucially, our GAT serves as an additive *compensator* to the backstepping control law rather than as a decision or policy generator: its output $\boldsymbol{u}_{\text{gat}}$ augments an already-stable nominal control, so the closed-loop system inherits the backstepping-layer stability certificate even when the attention weights are still adapting.

We regard each UAV as a node in the graph network, and we write the node set as

$$\mathcal{V} = \{1, 2, \ldots, N\}$$

Let R_{comm} denote the communication radius between UAVs. Then

$$a_{ij}(t) = \begin{cases} 1, & i \neq j,\ \|\boldsymbol{p}_i(t) - \boldsymbol{p}_j(t)\| \leq R_{\text{comm}}, \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

The distance-based adjacency matrix $\boldsymbol{A}(t) = [a_{ij}(t)]_{i,j=1}^{N}$ collects the communication relationships among UAVs, where the (i, j)-th entry is exactly the element $a_{ij}(t)$ defined above.

Based on the adjacency matrix, the neighbor set of the i-th UAV follows

$$\mathcal{N}_i(t) = \left\{ j \in \mathcal{V} \setminus \{i\} \,\middle|\, \|\boldsymbol{p}_i(t) - \boldsymbol{p}_j(t)\| \leq R_{\text{comm}} \right\} \tag{8}$$

Therefore, the time-varying graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ describes the UAV swarm, and the edge set $\mathcal{E}(t) = \{(i, j) : a_{ij}(t) = 1\}$ evolves with the relative positions of the UAVs. When node i has no neighbors within communication range, the softmax

normalization becomes undefined because its denominator vanishes. In that case, we set the aggregated feature $\boldsymbol{H}_{\mathrm{agg},i}(t)$ (defined in Section 2.3.2) to a zero vector, i.e.,

$$\mathcal{N}_i(t) = \emptyset \implies \boldsymbol{H}_{\mathrm{agg},i}(t) = \mathbf{0} \tag{9}$$

Based on this graph, the risk-aware attention aggregation is constructed in the following subsection.

2.3 Design of UAV Control Law

2.3.1 Nominal Sliding-Surface Tracking Control Law $\boldsymbol{u}_{\mathbf{track}}$

We introduce a first-order sliding surface to convert the second-order tracking problem into a first-order stabilization problem:

$$\boldsymbol{s}_i = \boldsymbol{e}_{v,i} + c\boldsymbol{e}_{p,i} \tag{10}$$

Here $c > 0$ denotes the weight gain of the sliding surface. When $\boldsymbol{s}_i = \mathbf{0}$, the error dynamics satisfy $\dot{\boldsymbol{e}}_{p,i} + c\boldsymbol{e}_{p,i} = \mathbf{0}$. The position error then converges exponentially with time constant $1/c$. Consequently, the control objective reduces to driving $\boldsymbol{s}_i \to \mathbf{0}$. Taking the time derivative of $\boldsymbol{s}_i$ along the system trajectory yields

$$\dot{\boldsymbol{s}}_i = \dot{\boldsymbol{v}}_i - \dot{\boldsymbol{v}}_{d,i} + c\dot{\boldsymbol{e}}_{p,i} = \big(\boldsymbol{u}_i + \boldsymbol{\delta}_i\big) - \boldsymbol{a}_L + c\boldsymbol{e}_{v,i} \tag{11}$$

For the nominal system without disturbances ($\boldsymbol{\delta}_i = \mathbf{0}$), we enforce $\dot{\boldsymbol{s}}_i = -K_A f_{\mathrm{obs},i} K_s \boldsymbol{s}_i$. The obstacle-distance gain $f_{\mathrm{obs},i}$ shapes the convergence rate according to the obstacle risk. After algebraic rearrangement, we obtain the tracking control law

$$\boldsymbol{u}_{\mathrm{track},i} = \boldsymbol{a}_L - c\boldsymbol{e}_{v,i} - \big(K_A f_{\mathrm{obs},i}\big) K_s \boldsymbol{s}_i \tag{12}$$

In this expression, $\boldsymbol{a}_L$ serves as the acceleration feedforward of the actual leader to achieve precise following. The term $-c\boldsymbol{e}_{v,i}$ acts as a damping term that regulates error convergence and suppresses velocity oscillations during tracking. The term $-K_A f_{\mathrm{obs},i} K_s \boldsymbol{s}_i$ represents the obstacle-modulated sliding reaching law, where K_A is the sliding gain coefficient and K_s is the sliding-surface gain.

The factor $f_{\mathrm{obs},i} \in (0, 1]$ is an obstacle-distance authority factor implemented from the minimum obstacle clearance $d_{\mathrm{obs},i}$:

$$\bar{d}_{\mathrm{obs},i} = \min\big(d_{\mathrm{obs},i}, d_{\mathrm{obs}}^{\mathrm{inf}}\big), \qquad f_{\mathrm{obs},i} = \frac{1}{2}\left[1 + \tanh\left(2\big(\bar{d}_{\mathrm{obs},i} - d_{\mathrm{obs}}^{\mathrm{safe}}\big)\right)\right]. \tag{13}$$

The influence radius $d_{\mathrm{obs}}^{\mathrm{inf}}$ only caps the distance used in this factor. In the certified safe operating region $d_{\mathrm{obs},i} \geq d_{\mathrm{obs}}^{\mathrm{safe}}$, one has $f_{\mathrm{obs},i} \geq 0.5$; far from obstacles, $f_{\mathrm{obs},i} \to 1$. The adaptive-gain ablation sets this factor to one.

2.3.2 Graph Attention Adaptive Compensation Term $\boldsymbol{u}_{\mathbf{gat}}$

When time-varying and impulsive disturbances act on the UAV, the nominal sliding-surface approach cannot compensate for $\boldsymbol{\delta}_i(t)$ in time, producing steady-state tracking errors. We use a graph attention network (GAT) to identify disturbance residuals

online and generate the compensation acceleration $\boldsymbol{u}_{\mathrm{gat}}$, so the closed-loop system adapts to unmodeled dynamics.

We construct the local feature vector of follower i by concatenating the position error, velocity error, external obstacle risk, inter-UAV collision risk, altitude risk, and vertical velocity:

$$\boldsymbol{\varphi}_i = \left[\boldsymbol{e}_{p,i}^{\top}, \boldsymbol{e}_{v,i}^{\top}, r_{\mathrm{obs},i}, r_{\mathrm{col},i}, z_{\mathrm{risk},i}, v_{z,i}\right]^{\top} \in \mathbb{R}^{10}, \tag{14}$$

where the three risk terms, which are used later in the GAT edge score q_{ij}, are

$$\begin{cases} r_{\mathrm{obs},i} = \displaystyle\sum_{q} \max\left(0, \frac{d_{\mathrm{obs}}^{\mathrm{inf}} - d_{\mathrm{obs},i,q}}{d_{\mathrm{obs}}^{\mathrm{inf}}}\right), \\ r_{\mathrm{col},i} = \displaystyle\sum_{j \neq i} \max\left(0, \frac{d_{\mathrm{uav}}^{\mathrm{inf}} - d_{ij}}{d_{\mathrm{uav}}^{\mathrm{inf}}}\right), \\ z_{\mathrm{risk},i} = \max(0, z_{\min} - p_{i,z}) + \max(0, p_{i,z} - z_{\max}), \end{cases} \tag{15}$$

with $d_{ij} = \|\boldsymbol{p}_i - \boldsymbol{p}_j\|$ the inter-UAV distance and $p_{i,z}$ the altitude of UAV i.

The graph attention mechanism computes an edge score q_{ij} between connected UAVs as

$$q_{ij} = \boldsymbol{a}_{\mathrm{att}}^{\top} \cdot \tfrac{1}{2}\big(\boldsymbol{\varphi}_i + \boldsymbol{\varphi}_j\big) + \lambda\big(r_{\mathrm{obs},j} + r_{\mathrm{col},j} + z_{\mathrm{risk},j}\big) + \lambda_v \|\boldsymbol{v}_i - \boldsymbol{v}_j\|, \tag{16}$$

where $\boldsymbol{a}_{\mathrm{att}} \in \mathbb{R}^{10}$ is a fixed attention-scoring vector, λ is the tuned risk-bias weight, and $\lambda_v = 0.6$ is a fixed relative-velocity weight outside the optimized gain vector. Negative scores are leaky-scaled before normalization,

$$\psi(q) = \max(q, 0.2q), \qquad \alpha_{ij} = \frac{\exp(\psi(q_{ij}))}{\sum_{k \in \mathcal{N}_i} \exp(\psi(q_{ik}))}. \tag{17}$$

The fixed scoring vector used in the simulations is $\boldsymbol{a}_{\mathrm{att}} = [0.20, 0.20, 0.18, 0.12, 0.12, 0.10, 0.18, 0.18, 0.10, 0.08]^{\top}$. The softmax normalization of the attention scores is illustrated in Fig. 1.

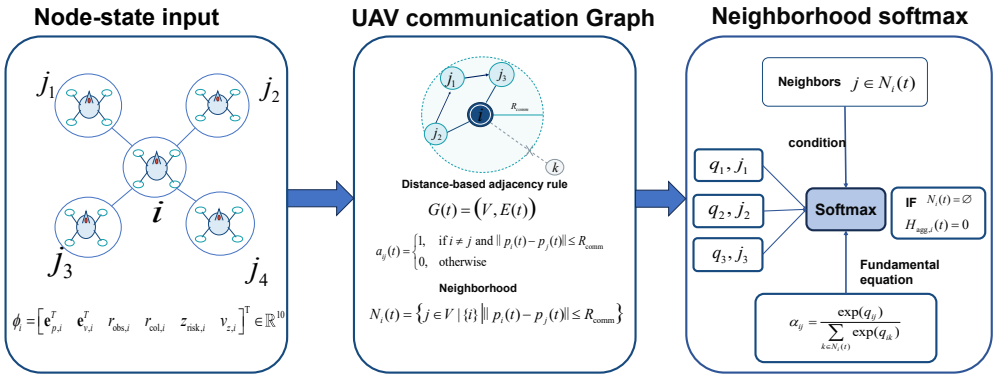


Fig. 1 Softmax normalization of attention scores.

With the normalized attention weights α_{ij}, the aggregated neighbor feature is computed as follows; Fig. 2 illustrates this weighted aggregation from the neighbors to UAV i:

$$\boldsymbol{H}_{\mathrm{agg},i} = \sum_{j\in\mathcal{N}_i} \alpha_{ij}\boldsymbol{\varphi}_j \in \mathbb{R}^{10}. \tag{18}$$

The aggregation reinforces compensation toward the most critical neighbors and, because followers share airspace and experience spatially correlated disturbances, $\boldsymbol{H}_{\mathrm{agg},i}$ provides an indirect yet timely estimate of the unmodeled disturbance residual acting on the current UAV.

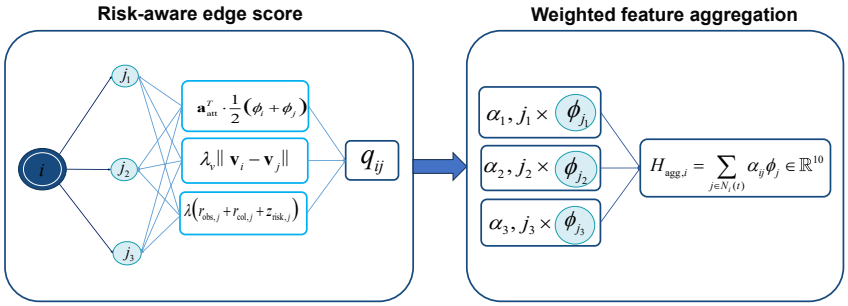


Fig. 2 Weighted feature aggregation from neighbors to UAV i.

The compensation term is constructed via a learnable weight matrix $\widehat{\boldsymbol{W}} \in \mathbb{R}^{10\times 3}$ as

$$\begin{cases} \widehat{\boldsymbol{\Delta}}_i = \widehat{\boldsymbol{W}}^{\top}\boldsymbol{H}_{\mathrm{agg},i} \in \mathbb{R}^3, \\ \boldsymbol{u}_{\mathrm{gat},i} = -g_{\mathrm{gate},i}K_G\widehat{\boldsymbol{\Delta}}_i. \end{cases} \tag{19}$$

where K_G is the compensation gain and the safety gating factor is

$$g_{\mathrm{gate},i} = \tfrac{1}{2}\left[1+\tanh\left((d_{\mathrm{obs},i}-d_{\mathrm{obs}}^{\mathrm{safe}})\cdot 3\right)\right] \in (0,1], \tag{20}$$

which drops below 0.5 as the UAV approaches obstacles and suppresses $\boldsymbol{u}_{\mathrm{gat}}$ when collision risk is high.

The weight matrix $\widehat{\boldsymbol{W}}$ is updated online at the simulation step Δt and projected onto a Frobenius-norm ball:

$$\begin{cases} \dot{\widehat{\boldsymbol{W}}} = \gamma \sum_{i=1}^{N} \boldsymbol{H}_{\mathrm{agg},i} \boldsymbol{s}_i^\top, \\ \widehat{\boldsymbol{W}}^{+} = \Pi_{\widehat{W}_m}\left(\widehat{\boldsymbol{W}} + \gamma\, \Delta t \sum_i \boldsymbol{H}_{\mathrm{agg},i} \boldsymbol{s}_i^\top \right), \end{cases} \tag{21}$$

where $\widehat{W}_m = 1.5$ and $\Pi_{\widehat{W}_m}$ rescales the matrix if $\|\widehat{\boldsymbol{W}}\|_F > \widehat{W}_m$, preventing drift when the compensation channel is gated off.

2.3.3 Artificial Potential Field Obstacle Avoidance Term $\boldsymbol{u}_{\mathbf{obs}}$

The obstacle-avoidance channel works on top of the GAT-compensated tracking law. Each obstacle q is represented by a type-specific signed-distance function $h_q(\boldsymbol{p})$, negative inside the obstacle and non-negative outside. The implementation uses spheres, axis-aligned boxes, cylinders, ellipsoids, and a torus. The clearance and outward direction are

$$\mathrm{clearance}_{i,q} = h_q(\boldsymbol{p}_i), \qquad \hat{\boldsymbol{a}}_{i,q} = \nabla h_q(\boldsymbol{p}_i) / \|\nabla h_q(\boldsymbol{p}_i)\| \tag{22}$$

away from nonsmooth boundary faces, where the code uses the corresponding type-specific outward face normal. Two representative cases are as follows.

For spherical obstacles:

$$\begin{cases} \mathrm{clearance} = \|\boldsymbol{p}_i - \boldsymbol{c}_q\| - r_q, \\ \hat{\boldsymbol{a}} = \dfrac{\boldsymbol{p}_i - \boldsymbol{c}_q}{\|\boldsymbol{p}_i - \boldsymbol{c}_q\|} \end{cases} \tag{23}$$

For box-shaped obstacles:

$$\begin{cases} \mathrm{clearance} = \big\|\boldsymbol{p}_i - \mathrm{closest}(\boldsymbol{p}_i, \mathrm{box})\big\|, \\ \hat{\boldsymbol{a}} = \dfrac{\boldsymbol{p}_i - \mathrm{closest}(\boldsymbol{p}_i, \mathrm{box})}{\big\|\boldsymbol{p}_i - \mathrm{closest}(\boldsymbol{p}_i, \mathrm{box})\big\|} \end{cases} \tag{24}$$

Cylinders, ellipsoids, and the torus use the same sign convention through their corresponding closest-point or level-set constructions. The minimum clearance of follower i to all obstacles reads

$$d_{\mathrm{obs},i} = \min_q \mathrm{clearance}(\boldsymbol{p}_i, O_q) \tag{25}$$

For each obstacle q, a repulsive force is activated when clearance $< d_{\mathrm{obs}}^{\mathrm{inf}}$:

$$\boldsymbol{rep}_{i,q} = w_z \eta_o \, \frac{\frac{1}{d_{\mathrm{eff}}} - \frac{1}{d_{\mathrm{obs}}^{\mathrm{inf}}}}{d_{\mathrm{eff}}^2} \, \hat{\boldsymbol{a}} \tag{26}$$

This expression gives the repulsive force from a single obstacle. The term $d_{\text{eff}} = \max(\text{clearance}, 0.08)$ avoids singularity. The weighting term

$$w_z = \frac{1}{2}\left[1 - \tanh\left(4(\text{clearance} - d_{\text{obs}}^{\text{safe}})\right)\right] \tag{27}$$

acts as a smooth switching function that adjusts the magnitude of the obstacle-avoidance control. When the UAV stays in the safe region, the switching weight approaches 0. It then weakens the repulsive effect automatically and lets the GAT compensation term restore the formation trajectory. When the UAV approaches the safety threshold, the switching weight rises smoothly to 1 and prioritizes the repulsive force over formation compensation. η_o denotes the external repulsive force coefficient. Note that w_z is complementary to the gating factor $g_{\text{gate},i}$ in Section 2.3.2: as the UAV approaches an obstacle, w_z rises toward 1 while $g_{\text{gate},i}$ falls toward 0, so the control authority shifts from the learned compensation channel to the explicit avoidance channel.

The repulsive force has strong nonlinearity proportional to $1/d^2$. Therefore, its magnitude increases rapidly as the UAV approaches obstacles. Summing the repulsive forces over all obstacles gives the total external repulsive force

$$\boldsymbol{rep}_{\text{Ext},i} = \sum_q \boldsymbol{rep}_{i,q} \tag{28}$$

UAVs often encounter dynamic obstacles during flight. Using only repulsive forces from instantaneous positions leads to lag effects. To handle this, the repulsive force is also evaluated at the predicted UAV position $\boldsymbol{p}_i(t) + \boldsymbol{v}_i(t)t_{\text{pred}}$ and at the predicted obstacle position at $t + t_{\text{pred}}$, with $t_{\text{pred}} = 0.5\,\text{s}$. For the periodic dynamic obstacles used here, the predicted center is obtained from the known sinusoidal obstacle-motion model; static obstacles remain fixed. The final repulsive force is the convex combination

$$\boldsymbol{rep}_{\text{Ext,smooth},i} = 0.7\boldsymbol{rep}_{\text{Ext},i} + 0.3\boldsymbol{rep}_{\text{Ext,pred},i} \tag{29}$$

This formulation removes the lag effect that arises when using current obstacle information to avoid future UAV positions.

The obstacle channel reuses the authority factor $f_{\text{obs},i}$ in Section 2.3.1. As clearance decreases, $f_{\text{obs},i}$ reduces the tracking reaching gain while $2 - f_{\text{obs},i}$ increases the repulsion gain. Thus the final obstacle-avoidance control term is

$$\boldsymbol{u}_{\text{obs},i} = K_R\left(2 - f_{\text{obs},i}\right)\boldsymbol{rep}_{\text{Ext,smooth},i} \tag{30}$$

This pair $(f_{\text{obs},i}, 2 - f_{\text{obs},i})$ gives a smooth authority exchange between tracking and explicit obstacle avoidance.

2.3.4 Inter-UAV Collision Avoidance Term $\boldsymbol{u}_{\text{col}}$

$$\boldsymbol{rep}_{ij}^{\text{col}} = \mathbb{I}[d_{ij} < d_{\text{uav}}^{\text{inf}}]\, w_s^u \eta_u \frac{\frac{1}{d_{\text{eff}}^u} - \frac{1}{d_{\text{uav}}^{\text{inf}}}}{\left(d_{\text{eff}}^u\right)^2} \frac{\boldsymbol{p}_i - \boldsymbol{p}_j}{d_{ij}} \tag{31}$$

where

$$w_s^u = \frac{1}{2}\left[1 - \tanh\left(4 \cdot \left(d_{ij} - d_{\text{uav}}^{\text{safe}}\right)\right)\right] \tag{32}$$

denotes the smooth switching function for inter-UAV collision avoidance, and its output lies in $(0, 1)$. The term $d_{\text{eff}}^u = \max\left(d_{ij}, 0.82 d_{\text{uav}}^{\text{safe}}\right)$ acts as an effective distance truncation to avoid singularities, where 0.82 is a fixed truncation coefficient, and η_u denotes the internal repulsive coefficient between UAVs. The total inter-UAV repulsive force reads

$$\boldsymbol{rep}_{\text{Int},i} = \sum_{j \neq i} \boldsymbol{rep}_{ij}^{\text{col}} \tag{33}$$

We formulate the inter-UAV collision-avoidance control term as

$$\boldsymbol{u}_{\text{col},i} = K_I \boldsymbol{rep}_{\text{Int},i} \tag{34}$$

where K_I denotes the gain for inter-UAV collision avoidance.

2.3.5 Altitude Soft Constraint Term

When the UAV swarm traverses obstacle regions, the flight altitude of each UAV is constrained within $[z_{\min}, z_{\max}]$. We measure the altitude $p_{i,z}$ from the take-off plane, where $z = 0$, and the interval $[z_{\min}, z_{\max}] = [-1.40, 1.90]$ m covers the take-off/climb and descent/landing phases. The core formulas read as follows:

$$\begin{cases} u_{\text{push up}} = G_z \, S\big((z_{\min} + 0.15) - p_{i,z}; 15, 0\big) \\ u_{\text{push down}} = G_z \, S\big(p_{i,z} - (z_{\max} - 0.15); 15, 0\big) \\ \boldsymbol{u}_{\text{alt},i} = \big[0,\ 0,\ u_{\text{push up}} - u_{\text{push down}}\big] \end{cases} \tag{35}$$

Here the sigmoid function takes the form

$$S(z; k, z_0) = \frac{1}{1 + \exp\big(-k(z - z_0)\big)} \tag{36}$$

$G_z = 1.5|z_{\min}|$ denotes the amplitude of the upward and downward altitude adjustment. When the altitude falls below $z_{\min} + 0.15$, the controller activates an upward push; when it exceeds $z_{\max} - 0.15$, a downward push activates.

2.3.6 Safety-Conflict Projection and Input Saturation

Before the five channels are summed, the implementation applies a conflict projection when the nominal tracking command points against the obstacle-repulsion command near an obstacle. Let

$$\hat{\boldsymbol{u}}_{\text{obs},i} = \frac{\boldsymbol{u}_{\text{obs},i}}{\|\boldsymbol{u}_{\text{obs},i}\|}, \qquad \rho_i^{\text{ACP}} = \frac{1}{2}\left[1 - \tanh\left(8(d_{\text{obs},i} - 0.30)\right)\right]. \tag{37}$$

The projection is active when $d_{\text{obs},i} < 0.55\,\text{m}$ and

$$\frac{\boldsymbol{u}_{\text{track},i}^{\top} \boldsymbol{u}_{\text{obs},i}}{\|\boldsymbol{u}_{\text{track},i}\| \, \|\boldsymbol{u}_{\text{obs},i}\|} < -0.20.$$

Then,

$$\boldsymbol{u}_{\mathrm{track},i}^{\mathrm{ACP}} = \boldsymbol{u}_{\mathrm{track},i} - \rho_i^{\mathrm{ACP}} \left(\boldsymbol{u}_{\mathrm{track},i}^{\top} \hat{\boldsymbol{u}}_{\mathrm{obs},i}\right) \hat{\boldsymbol{u}}_{\mathrm{obs},i}; \tag{38}$$

otherwise $\boldsymbol{u}_{\mathrm{track},i}^{\mathrm{ACP}} = \boldsymbol{u}_{\mathrm{track},i}$. This removes the component of the tracking command that directly opposes obstacle repulsion while retaining the tangential tracking component.

2.3.7 Complete Control Law Formulation

In summary, the raw command for the i-th follower UAV reads

$$\begin{aligned} \boldsymbol{u}_i^{\mathrm{raw}} = \; & \underbrace{\boldsymbol{u}_{\mathrm{track},i}^{\mathrm{ACP}}}_{\boldsymbol{u}_{\mathrm{track}}} + \underbrace{\left(-g_{\mathrm{gate},i} K_G \widehat{\boldsymbol{W}}^{\top} \boldsymbol{H}_{\mathrm{agg},i}\right)}_{\boldsymbol{u}_{\mathrm{gat}}} \\ & + \underbrace{K_R \left(2 - f_{\mathrm{obs},i}\right) \boldsymbol{rep}_{\mathrm{Ext,smooth},i}}_{\boldsymbol{u}_{\mathrm{obs}}} \\ & + \underbrace{K_I \boldsymbol{rep}_{\mathrm{Int},i}}_{\boldsymbol{u}_{\mathrm{col}}} + \underbrace{\boldsymbol{u}_{\mathrm{alt},i}}_{\boldsymbol{u}_{\mathrm{alt}}} \end{aligned} \tag{39}$$

The applied acceleration is

$$\boldsymbol{u}_i = \begin{cases} \boldsymbol{u}_i^{\mathrm{raw}}, & \|\boldsymbol{u}_i^{\mathrm{raw}}\| \le U_{\max}, \\ \dfrac{\boldsymbol{u}_i^{\mathrm{raw}}}{\|\boldsymbol{u}_i^{\mathrm{raw}}\|} U_{\max} \tanh\left(\dfrac{\|\boldsymbol{u}_i^{\mathrm{raw}}\|}{U_{\max}}\right), & \|\boldsymbol{u}_i^{\mathrm{raw}}\| > U_{\max}. \end{cases} \tag{40}$$

This bounded saturation map retains the command direction. Fig. 3 summarizes the five parallel channels; the conflict projection and saturation are implementation-level safety post-processors.

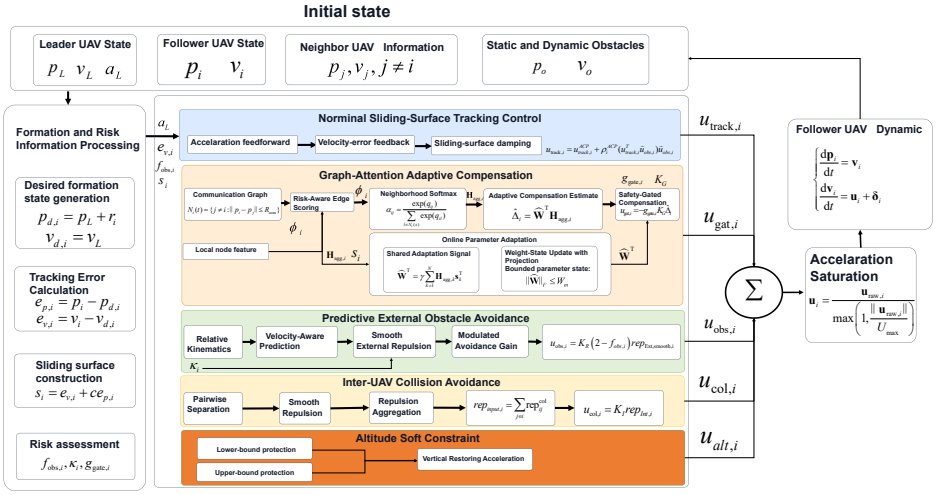


Fig. 3 Five-channel GAT-based control law before implementation-level conflict projection and saturation.

2.3.8 A Unified Safety-Margin Perspective

Although the five channels were derived channel by channel, they are coordinated by a single design variable: the *safety margin* of follower i with respect to obstacles,

$$m_{\mathrm{obs},i}(t) \;\triangleq\; d_{\mathrm{obs},i}(t) - d_{\mathrm{obs}}^{\mathrm{safe}}, \tag{41}$$

and, symmetrically, the inter-UAV margin $m_{\mathrm{uav},ij} \triangleq d_{ij} - d_{\mathrm{uav}}^{\mathrm{safe}}$ between followers i and j. The margin is positive in the safe region, vanishes at the hard threshold, and becomes negative upon violation; every safety mechanism of the proposed law is either a function of $m_{\mathrm{obs},i}$ or an operator acting on it. Table 1 summarizes how the safety-oriented design degrees of freedom of the control law organize into four layers, each regulating one aspect of the margin.

Two features of this layered design deserve emphasis. First, the handover among the tracking, compensation and avoidance channels is smooth by construction—the structural origin of the low-chatter command histories and the absence of sustained saturation reported in Section 3. Second, the design is *margin-regulating* rather than margin-guaranteeing: the Lyapunov analysis of Section 2.3.9 certifies bounded tracking error, whereas the persistence of a positive margin rests on the closed-loop simulations rather than on a formal invariance argument; the inter-UAV margin, in particular, is regulated only indirectly through the formation-tracking layer and the repulsive term, which is the identified limitation behind the sub-unitary joint success rates in dense scenes. The ablation study of Section 3.5 provides the experimental counterpart of this table: the safety-critical backbone modules (No-Obs, No-Predict, No-Track)

Table 1 Layered safety-margin design of the proposed control law.

Layer	Design variables	Values	Role in margin regulation
Threshold	$d_{\text{obs}}^{\text{safe}}$, $d_{\text{obs}}^{\text{inf}}$; $d_{\text{uav}}^{\text{safe}}$, $d_{\text{uav}}^{\text{inf}}$	0.22/1.45 m; 0.65/1.25 m	Defines the margin domain: 1.23 m (obstacle) and 0.60 m (inter-UAV) buffer bands between the warning radius and the hard threshold $m = 0$
Switching	tanh slopes of $g_{\text{gate},i}$, w_z, w_s^u, and $f_{\text{obs},i}$	3, 4, 4, 2	Schedules the margin-for-authority exchange: the complementary pair $(g_{\text{gate},i}, w_z)$ hands authority from the compensation channel to the avoidance channel as $m_{\text{obs},i} \to 0$
Action	t_{pred}; convex weight 0.7/0.3; authority pair $(f_{\text{obs},i}, 2 - f_{\text{obs},i})$; ACP threshold/cosine/slope; truncations d_{eff}, d_{eff}^u	0.5 s; 0.7/0.3; [1, 2]; 0.55 m, −0.20, 8; 0.08 m, $0.82 d_{\text{uav}}^{\text{safe}}$	Extends the margin in time and space: prediction converts future margin into present reaction, the authority pair and ACP resolve tracking–avoidance conflict, and truncations keep the command bounded
Optimization	$\eta_o, \eta_u, K_R, K_I \in \boldsymbol{\theta}$; Pen_{safe} structure	CMA-ES tuned; feasibility-first	Promotes the margin from a constraint to an objective: the CMA-ES objective maps violations to a larger fitness than any feasible candidate and mildly rewards wider margins (Section 2.4.1)

correspond to the Action and nominal-tracking layers, confirming that the layers acting most directly on the margin are also the ones whose removal breaks it.

2.3.9 Stability Analysis

The following assumptions are made for the stability analysis.

Assumption 1 (Leader state) *The leader trajectory* $(\boldsymbol{p}_L, \boldsymbol{v}_L, \boldsymbol{a}_L)$ *is known, bounded and sufficiently smooth with* $\|\boldsymbol{v}_{L,xy}\| \le v_{\max}$*, and is broadcast to all followers via a reliable channel independent of* $\mathcal{G}(t)$*.*

Assumption 2 (Disturbance) $\|\boldsymbol{\delta}_i(t)\| \le \delta_{\max}$ *with* $\delta_{\max} = 0.5\,\text{m/s}^2$*, and* $\boldsymbol{\delta}_i = g_{\text{gate},i} K_G \boldsymbol{W}^{*\top} \boldsymbol{H}_{\text{agg},i} + \boldsymbol{\eta}_i$ *with* $\|\boldsymbol{\eta}_i\| \le \eta_{\max}$ *and* $\|\boldsymbol{W}^*\|_F \le W_m$*.*

The decomposition in Assumption 2 is a standard linearly parameterized (matching-condition) assumption in adaptive control: the component of the lumped disturbance that lies in the span of the GAT regressor $\boldsymbol{H}_{\text{agg},i}$ is captured by the ideal weight matrix $\boldsymbol{W}^*$, while any structural mismatch—e.g., disturbance components outside the regressor span or time variation of $\boldsymbol{W}^*$ itself—is absorbed into the residual $\boldsymbol{\eta}_i$. This assumption does not require the physical disturbance to actually possess this structure; a violation only enlarges the residual bound $\eta_{\max}$ and hence the UUB radius through $b = \eta_{\max} + \zeta_{\max}$, without invalidating the boundedness result of Section 2.3.9.

Assumption 3 (Measurability) *The states* $\boldsymbol{p}_i, \boldsymbol{v}_i$*, the leader state, neighbor positions, and obstacle positions/velocities are measurable.*

Assumption 4 (Boundedness) $\|\boldsymbol{H}_{\mathrm{agg},i}\| \le H_m$; *the projection confines* $\|\widehat{\boldsymbol{W}}\|_F \le \widehat{W}_m$, *hence* $\|\widetilde{\boldsymbol{W}}\|_F \le \widetilde{W}_m \triangleq \widehat{W}_m + W_m$; *and* $\|\boldsymbol{\zeta}_i\| \le \zeta_{\max}$ *for* $\boldsymbol{\zeta}_i = \boldsymbol{u}_{\mathrm{obs},i} + \boldsymbol{u}_{\mathrm{col},i} + \boldsymbol{u}_{\mathrm{alt},i}$.

Assumption 5 (Input authority) $U_{\max}$ *is chosen so that saturation stays inactive in the operating region, where the UUB analysis applies directly.*

The following Lyapunov derivation concerns the smooth unprojected law with $\boldsymbol{u}^{\mathrm{ACP}}_{\mathrm{track},i} = \boldsymbol{u}_{\mathrm{track},i}$ and with the saturation map inactive. The conflict projection and saturation are bounded safety post-processors used by the implementation; their numerical effect is assessed in Section 3.

We establish uniform ultimate boundedness (UUB) of the closed-loop system by constructing a composite Lyapunov candidate that combines the sliding-surface energy with the GAT weight-estimation error. Let $\boldsymbol{s} = [\boldsymbol{s}_1^\top, \dots, \boldsymbol{s}_N^\top]^\top \in \mathbb{R}^{3N}$ denote the stacked sliding-surface vector and $\widetilde{\boldsymbol{W}} = \widehat{\boldsymbol{W}} - \boldsymbol{W}^* \in \mathbb{R}^{10\times 3}$ the parameter estimation error, whose Frobenius-norm bound is $\widetilde{W}_m \triangleq \widehat{W}_m + W_m$ (Assumption 4). We define

$$V(\boldsymbol{s}, \widetilde{\boldsymbol{W}}) = \frac{1}{2}\sum_{i=1}^{N} \boldsymbol{s}_i^\top \boldsymbol{s}_i + \frac{1}{2\gamma}\mathrm{tr}\Big(\widetilde{\boldsymbol{W}}^\top \widetilde{\boldsymbol{W}}\Big). \tag{42}$$

By construction, V is positive definite, radially unbounded, and C^1 in $(\boldsymbol{s}, \widetilde{\boldsymbol{W}})$.

Substituting the complete control law of Section 2.3.7 together with the weight-update law $\dot{\widehat{\boldsymbol{W}}} = \gamma \sum_{i=1}^{N} \boldsymbol{H}_{\mathrm{agg},i}\boldsymbol{s}_i^\top$ into $\dot{V} = \sum_i \boldsymbol{s}_i^\top \dot{\boldsymbol{s}}_i + \frac{1}{\gamma}\mathrm{tr}(\widetilde{\boldsymbol{W}}^\top \dot{\widehat{\boldsymbol{W}}})$ yields

$$\dot{V} = -\sum_{i=1}^{N} K_A f_{\mathrm{obs},i} K_s \|\boldsymbol{s}_i\|^2 + \sum_{i=1}^{N}\big(1 - g_{\mathrm{gate},i}K_G\big)\,\mathrm{tr}\big(\widetilde{\boldsymbol{W}}^\top \boldsymbol{H}_{\mathrm{agg},i}\boldsymbol{s}_i^\top\big) + \sum_{i=1}^{N} \boldsymbol{s}_i^\top\big(\boldsymbol{\eta}_i + \boldsymbol{\zeta}_i\big), \tag{43}$$

where $\boldsymbol{\zeta}_i \triangleq \boldsymbol{u}_{\mathrm{obs},i} + \boldsymbol{u}_{\mathrm{col},i} + \boldsymbol{u}_{\mathrm{alt},i}$ collects the auxiliary avoidance/constraint terms. The first term provides sliding reaching-law damping, the second couples the estimation error with the aggregated feature, and the third absorbs the residual disturbance and the avoidance reaction.

To convert this expression into a closed-form bound, we set $b \triangleq \eta_{\max} + \zeta_{\max}$ using Assumptions 2 and 4, and bound the trace cross-term through Cauchy–Schwarz as $\mathrm{tr}(\widetilde{\boldsymbol{W}}^\top \boldsymbol{H}_{\mathrm{agg},i}\boldsymbol{s}_i^\top) \le \|\widetilde{\boldsymbol{W}}\|_F\,\|\boldsymbol{H}_{\mathrm{agg},i}\|\,\|\boldsymbol{s}_i\| \le H_m\|\widetilde{\boldsymbol{W}}\|_F\|\boldsymbol{s}_i\|$. Together with the gate-amplitude bound $|1 - g_{\mathrm{gate},i}K_G| \le 1 + K_G$, this gives the coupling coefficient $\varrho \triangleq (1 + K_G)H_m$. Summing over all i and using the Cauchy–Schwarz bound $\sum_{i=1}^{N}\|\boldsymbol{s}_i\| \le \sqrt{N}\,\|\boldsymbol{s}\|$, the last two terms of Eq. (43) satisfy

$$\sum_{i=1}^{N} \varrho\,\|\widetilde{\boldsymbol{W}}\|_F\|\boldsymbol{s}_i\| + \sum_{i=1}^{N} b\,\|\boldsymbol{s}_i\| \;\le\; \big(\varrho\|\widetilde{\boldsymbol{W}}\|_F + b\big)\sqrt{N}\,\|\boldsymbol{s}\|. \tag{44}$$

Applying Young's inequality $ab \le \frac{\epsilon}{2}a^2 + \frac{1}{2\epsilon}b^2$ twice—first with $a = \|\boldsymbol{s}\|$, $b = \sqrt{N}\varrho\|\widetilde{\boldsymbol{W}}\|_F$, $\epsilon = \lambda_{\min}$, and then with $a = \|\boldsymbol{s}\|$, $b = \sqrt{N}\,b$, $\epsilon = \lambda_{\min}/2$—and using the

projection bound $\|\widetilde{\boldsymbol{W}}\|_F \le \widetilde{W}_m$ from Assumption 4 to convert the weight-error factor into a constant, we obtain

$$\dot{V} \;\le\; -\frac{\lambda_{\min}}{4}\|\boldsymbol{s}\|^2 + c^*, \qquad c^* \;\triangleq\; \frac{N\varrho^2\widetilde{W}_m^2}{2\lambda_{\min}} + \frac{N}{\lambda_{\min}}b^2, \tag{45}$$

where $\lambda_{\min} \triangleq K_A K_s \min_i f_{\mathrm{obs},i} > 0$. The first term is negative definite and dominates once $\|\boldsymbol{s}\|$ exceeds the corresponding threshold, whereas c^* is a constant that determines the UUB radius. Two remarks are in order. First, the boundedness of $\widetilde{\boldsymbol{W}}$ is not established by the Lyapunov decrease itself but by the projection operator in the update law (21), which confines $\|\widehat{\boldsymbol{W}}\|_F \le \widehat{W}_m$ so that $\|\widetilde{\boldsymbol{W}}\|_F \le \widetilde{W}_m$ holds at all times; the $\widetilde{\boldsymbol{W}}$-block of V therefore records a bounded energy and plays no role in the decrease condition. Second, the adaptation term cancels the $g_{\mathrm{gate},i}K_G$-portion of the cross-term exactly (as evidenced in Eq. (43)), and only the residual mismatch $(1 - g_{\mathrm{gate},i}K_G)$ feeds c^* through ϱ.

From Eq. (45), the inequality $\dot{V} < 0$ holds whenever

$$\|\boldsymbol{s}\| \;>\; \frac{2\sqrt{c^*}}{\sqrt{\lambda_{\min}}} \;\triangleq\; \mathcal{R}_s, \tag{46}$$

so V decreases monotonically whenever the stacked sliding surface leaves the ball of radius $\mathcal{R}_s$, and every closed-loop trajectory enters the compact set

$$\Omega = \big\{\,(\boldsymbol{s}, \widetilde{\boldsymbol{W}}) \;:\; V \le V_\infty\,\big\}, \qquad V_\infty \;\triangleq\; \frac{2c^*}{\lambda_{\min}} + \frac{\widetilde{W}_m^2}{2\gamma} = \frac{N\varrho^2\widetilde{W}_m^2 + 2Nb^2}{\lambda_{\min}^2} + \frac{\widetilde{W}_m^2}{2\gamma}, \tag{47}$$

in finite time and remains there. Indeed, since $\frac{\lambda_{\min}}{4}\|\boldsymbol{s}\|^2 \ge \frac{\lambda_{\min}}{2}\big(V - \frac{1}{2\gamma}\|\widetilde{\boldsymbol{W}}\|_F^2\big) \ge \frac{\lambda_{\min}}{2}V - \frac{\lambda_{\min}\widetilde{W}_m^2}{4\gamma}$, Eq. (45) and the comparison lemma give

$$V(t) \;\le\; V(0)\,e^{-\lambda_{\min}t/2} + \Big(\frac{2c^*}{\lambda_{\min}} + \frac{\widetilde{W}_m^2}{2\gamma}\Big)\Big(1 - e^{-\lambda_{\min}t/2}\Big), \tag{48}$$

so V converges exponentially to the compact, invariant level set $V \le V_\infty$ of Eq. (47). In terms of the original state, the sliding surface satisfies $\|\boldsymbol{s}_i\| \le \|\boldsymbol{s}\| \le \sqrt{2V_\infty}$ uniformly for all i and all $t \ge \mathcal{T} = \frac{2}{\lambda_{\min}}\ln\frac{V(0)}{V_\infty}$ (for $V(0) > V_\infty$; otherwise $\mathcal{T} = 0$), and the GAT weight estimate stays in the projection set with $\|\widetilde{\boldsymbol{W}}\|_F \le \widetilde{W}_m$ for all $t \ge 0$. This bound tightens monotonically with K_A and K_s, which is consistent with the offline gain tuning in Section 2.4.

From the sliding-surface definition $\boldsymbol{s}_i = \boldsymbol{e}_{v,i} + c\boldsymbol{e}_{p,i}$, the position error dynamics obey $\dot{\boldsymbol{e}}_{p,i} = -c\boldsymbol{e}_{p,i} + \boldsymbol{s}_i$, i.e., a stable linear system driven by the bounded input $\boldsymbol{s}_i$. For $c > 0$,

$$\|\boldsymbol{e}_{p,i}(t)\| \;\le\; \|\boldsymbol{e}_{p,i}(0)\|\,e^{-ct} + \frac{1}{c}\sup_{\tau\ge 0}\|\boldsymbol{s}_i(\tau)\|, \tag{49}$$

so $\boldsymbol{e}_{p,i}$ inherits UUB from $\boldsymbol{s}_i$, with asymptotic envelope $\|\boldsymbol{e}_{p,i}\|_\infty \le \sqrt{2V_\infty}/c$. Differentiating gives the velocity-error bound $\|\boldsymbol{e}_{v,i}\|_\infty \le \|\boldsymbol{s}_i\|_\infty + c\|\boldsymbol{e}_{p,i}\|_\infty \le (1+c^{-1})\sqrt{2V_\infty}$. Hence both tracking errors $\boldsymbol{e}_{p,i}$ and $\boldsymbol{e}_{v,i}$ are uniformly ultimately bounded with explicit envelopes.

Under the ideal conditions $\boldsymbol{\eta}_i \equiv \mathbf{0}$, $\boldsymbol{\zeta}_i \equiv \mathbf{0}$ and $g_{\mathrm{gate},i}K_G \equiv 1$, the residual term in Eq. (43) vanishes, leaving $\dot{V} \le -\lambda_{\min}\|\boldsymbol{s}\|^2$, so $\boldsymbol{s}_i \to \mathbf{0}$ asymptotically by LaSalle's invariance principle; the weight error $\widetilde{\boldsymbol{W}}$ converges to a constant matrix, and to zero under persistent excitation of $\boldsymbol{H}_{\mathrm{agg},i}$.

For the discrete-time implementation, the Euler forward discretization with step Δt approximates $\dot{V}$ by a Taylor expansion with a second-order remainder,

$$\Delta V_k = V\big(t_k+\Delta t\big) - V(t_k) \;\le\; \dot{V}(t_k)\,\Delta t + \tfrac{1}{2}\sup\|\ddot{V}\|\,\Delta t^2. \tag{50}$$

Since $V = \frac{1}{2}\sum_i \|\boldsymbol{s}_i\|^2 + \frac{1}{2\gamma}\|\widetilde{\boldsymbol{W}}\|_F^2$ and $\dot{\widehat{\boldsymbol{W}}}$ is bounded by Assumption 4, the dominant contribution to $\ddot{V}$ comes from $\sum_i \|\dot{\boldsymbol{s}}_i\|^2$. The reaching law gives $\|\dot{\boldsymbol{s}}_i\| \le K_A K_s\|\boldsymbol{s}_i\| + \varrho\|\widetilde{\boldsymbol{W}}\|_F + b$, and keeping the two state-proportional terms yields $\sup\|\ddot{V}\| \lesssim 2\big[(K_AK_s)^2 + \lambda_{\min}^2\big]\|\boldsymbol{s}\|^2$, where the $(K_AK_s)^2$ term reflects the fastest closed-loop convergence rate and the $\lambda_{\min}^2$ term reflects the residual decay rate. Requiring the second-order term $\big[(K_AK_s)^2 + \lambda_{\min}^2\big]\|\boldsymbol{s}\|^2\Delta t^2$ not to exceed the first-order decrease $\frac{\lambda_{\min}}{4}\|\boldsymbol{s}\|^2\Delta t$ of Eq. (45) gives the sufficient step-size condition

$$\Delta t \;\le\; \Delta t_{\max} \;\triangleq\; \frac{\lambda_{\min}}{4\big[(K_AK_s)^2 + \lambda_{\min}^2\big]}, \tag{51}$$

which guarantees $\Delta V_k \le -\frac{\lambda_{\min}}{4}\|\boldsymbol{s}_k\|^2\Delta t < 0$ outside the UUB set. Substituting the nominal baseline gains of Table 3 ($K_A = 1.0$, $K_s = 2.0$, $\min_i f_{\mathrm{obs},i} \ge 0.5$) gives $\lambda_{\min} = 1.0$ and $\Delta t_{\max} = \frac{1.0}{4\cdot(4.0+1.0)} = 0.05\,\mathrm{s}$, exactly matching the simulation step adopted throughout Section 3. For the CMA-ES tuned gains of Table 3 ($K_A = 1.4878$, $K_s = 1.2387$, hence $K_AK_s = 1.842$ and $\lambda_{\min} = 0.921$), the same formula yields $\Delta t_{\max} \approx 0.054\,\mathrm{s}$, so the adopted step satisfies the sufficient condition with an 8.5% margin. The tuned solution avoids the excessively large K_AK_s regime while retaining a fast admissible convergence rate. In the dense-scene validation, only 0.11% of samples reach the saturation envelope, so the unsaturated analysis is representative of almost the entire mission. When saturation does occur, the UUB bound no longer applies to those samples; nevertheless, the simulations remain numerically stable because the safety gate $g_{\mathrm{gate},i}$ smoothly transfers authority to the explicit avoidance channel. Likewise, when $\mathcal{N}_i = \emptyset$ the GAT compensator degrades to $\boldsymbol{H}_{\mathrm{agg},i} = \mathbf{0}$ (Eq. (9)), and the analysis reduces to the classical tracking-plus-avoidance closed loop, which is again UUB under the same assumptions.

2.4 Offline Gain Tuning by CMA-ES

The five control channels contain ten strongly coupled gains: repulsion strength trades against tracking bandwidth, GAT compensation must remain subordinate to safety, and mission-level safety indicators are non-smooth and seed-dependent. The gain

vector is therefore tuned by safety-constrained simulation optimization rather than manual search.

Table 2 Search bounds for the ten control gains used in offline CMA-ES tuning.

Parameters	K_A	K_R	K_I	K_s	c	η_o	η_u	K_G	λ	γ
Lower bound	0.3	0.3	0.3	0.5	0.3	0.02	0.01	0.1	0.1	0.005
Upper bound	2.5	4.0	2.5	4.5	2.0	0.25	0.20	1.8	4.0	0.12

2.4.1 Safety-Constrained Objective

For each closed-loop training simulation, the optimizer uses three normalized performance indicators: tracking RMSE J_e, control energy J_u, and control smoothness $J_{\Delta u}$. The scene-level performance cost is

$$\mathcal{J}_s = \omega_e \frac{J_{e,s}}{J_e^{\mathrm{ref}}} + \omega_u \frac{J_{u,s}}{J_u^{\mathrm{ref}}} + \omega_{\Delta u} \frac{J_{\Delta u,s}}{J_{\Delta u}^{\mathrm{ref}}}, \tag{52}$$

where

$$\omega_e + \omega_u + \omega_{\Delta u} = 1, \qquad (\omega_e, \omega_u, \omega_{\Delta u}) = (0.60, 0.20, 0.20).$$

These values were fixed in the experiment configuration before the formal CMA-ES runs: preliminary development checks indicated that a larger tracking weight with equal, moderate energy and smoothness regularization gave a stable safety–performance trade-off without over-regularizing the command. The weights were not optimized by CMA-ES and were never selected using the validation seeds. The reference scales $(J_e^{\mathrm{ref}}, J_u^{\mathrm{ref}}, J_{\Delta u}^{\mathrm{ref}}) = (0.368831, 58.6706, 116.9924)$ were calibrated separately on disjoint development instances and then frozen. Sliding-surface error, terminal error, and saturation rate are validation diagnostics rather than training-fitness terms. Safety is handled before performance. With

$$\nu_{\mathrm{obs}} = \max\left(0, \frac{d_{\mathrm{obs}}^{\mathrm{safe}} - d_{\mathrm{min}}^{\mathrm{obs}}}{d_{\mathrm{obs}}^{\mathrm{safe}}}\right), \quad \nu_{\mathrm{uav}} = \max\left(0, \frac{d_{\mathrm{uav}}^{\mathrm{safe}} - d_{\mathrm{min}}^{\mathrm{uav}}}{d_{\mathrm{uav}}^{\mathrm{safe}}}\right), \tag{53}$$

let $\nu_s = \nu_{\mathrm{obs}} + \nu_{\mathrm{uav}}$, and let $\nu_{\max} = \max_s \nu_s$ over all training scenes. A failed or incomplete run receives the frozen violation $\nu_s = 6$, larger than the violation of any complete run in these missions. If $S = 6$ and $J = S^{-1} \sum_s \mathcal{J}_s$, the scalar fitness is

$$F(\boldsymbol{z}) = \begin{cases} \dfrac{J}{1+J}, & \nu_{\max} \le 10^{-10} \text{ and all indicators are finite,} \\ 1 + \nu_{\max}, & \text{otherwise.} \end{cases} \tag{54}$$

Thus every feasible candidate has $F < 1$ and strictly dominates every infeasible candidate.

2.4.2 CMA-ES Search Model

CMA-ES models promising gain vectors by a multivariate normal distribution [26, 27]. Search is performed in $\boldsymbol{z} \in [0,1]^{10}$ and decoded as

$$\boldsymbol{\theta} = \boldsymbol{\ell} + (\bar{\boldsymbol{u}} - \boldsymbol{\ell}) \circ \boldsymbol{z}, \tag{55}$$

where $\boldsymbol{\ell}, \bar{\boldsymbol{u}}$ are the bounds in Table 2. At generation g,

$$\boldsymbol{z}_k^{(g+1)} = \mathcal{R}\Big[\boldsymbol{m}^{(g)} + \sigma^{(g)}\boldsymbol{B}^{(g)}\boldsymbol{n}_k\Big], \qquad \boldsymbol{n}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \tag{56}$$

for $k = 1, \ldots, \Lambda$, where $\boldsymbol{m}$ is the search center, σ is the global step, $\boldsymbol{B}\boldsymbol{B}^\top = \boldsymbol{C}$, and $\mathcal{R}$ reflects samples into $[0, 1]^{10}$. The covariance $\boldsymbol{C}$ stretches sampling along directions in which coupled gains have previously produced safe, low-cost missions. Initially $\boldsymbol{C} = \boldsymbol{I}$, $\sigma = 0.3$, $\Lambda = 4 + \lfloor 3 \ln 10 \rfloor = 10$, and $\mu = \lfloor \Lambda/2 \rfloor = 5$.

After closed-loop evaluation, the μ best candidates are selected. Their rank weights and effective selection mass are

$$w_i = \frac{\ln(\mu + 0.5) - \ln i}{\sum_{j=1}^{\mu}[\ln(\mu + 0.5) - \ln j]}, \qquad \mu_{\text{eff}} = \left(\sum_{i=1}^{\mu} w_i^2\right)^{-1}. \tag{57}$$

The center moves by weighted recombination,

$$\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \sum_{i=1}^{\mu} w_i(\boldsymbol{z}_{i:\Lambda}^{(g+1)} - \boldsymbol{m}^{(g)}). \tag{58}$$

Define $\boldsymbol{y}_i = (\boldsymbol{z}_{i:\Lambda}^{(g+1)} - \boldsymbol{m}^{(g)})/\sigma^{(g)}$ and $\boldsymbol{\delta}_m = (\boldsymbol{m}^{(g+1)} - \boldsymbol{m}^{(g)})/\sigma^{(g)}$. The conjugate and covariance evolution paths are

$$\boldsymbol{p}_\sigma^{(g+1)} = (1 - c_\sigma)\boldsymbol{p}_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\mu_{\text{eff}}}\,\boldsymbol{C}^{-1/2}\boldsymbol{\delta}_m, \tag{59}$$

$$\boldsymbol{p}_c^{(g+1)} = (1 - c_c)\boldsymbol{p}_c^{(g)} + h_\sigma\sqrt{c_c(2 - c_c)\mu_{\text{eff}}}\,\boldsymbol{\delta}_m, \tag{60}$$

where h_σ suppresses covariance learning after a too-rapid step-size change. The covariance update is

$$\boldsymbol{C}^{(g+1)} = (1 - c_1 - c_\mu)\boldsymbol{C}^{(g)} + c_1[\boldsymbol{p}_c\boldsymbol{p}_c^\top + (1 - h_\sigma)c_c(2 - c_c)\boldsymbol{C}^{(g)}] + c_\mu\sum_{i=1}^{\mu} w_i\boldsymbol{y}_i\boldsymbol{y}_i^\top, \tag{61}$$

and cumulative step-size adaptation gives

$$\sigma^{(g+1)} = \sigma^{(g)}\exp\left[\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma^{(g+1)}\|}{\chi_D} - 1\right)\right], \quad \chi_D \approx \sqrt{D}\left(1 - \frac{1}{4D} + \frac{1}{21D^2}\right). \tag{62}$$

If $\sigma < 10^{-6}$, the best fitness stagnates for 30 generations, $\text{cond}(\boldsymbol{C}) > 10^{12}$, or $\boldsymbol{C}$ becomes non-finite, a restart [28] doubles Λ up to 64, resets $\boldsymbol{C} = \boldsymbol{I}$ and both paths, sets $\sigma = 0.3$, and centers the new distribution at the historical best. This restart rule was enabled, but the selected full-dimensional run did not trigger a restart and retained $\Lambda = 10$. Algorithm 1 summarizes the offline procedure.

Algorithm 1: Safety-constrained restart-enabled CMA-ES for controller gain tuning

Input: bounds $[\boldsymbol{\ell}, \bar{\boldsymbol{u}}]$, training scenes, budget 816, hard-start mean
Output: feasible tuned gain vector $\boldsymbol{\theta}^{\star}$

Initialize $\Lambda = 10$, $\mu = 5$, $\sigma = 0.3$, $\boldsymbol{C} = \boldsymbol{I}$, $\boldsymbol{p}_c = \boldsymbol{p}_\sigma = \boldsymbol{0}$, and center $\boldsymbol{m}$ at the hard-start mean;
while *budget remains* **do**
  Sample Λ candidates $\boldsymbol{z}_k \sim \mathcal{N}(\boldsymbol{m}, \sigma^2 \boldsymbol{C})$ and reflect them into $[0,1]^{10}$;
  for *each candidate* $\boldsymbol{z}_k$ **do**
    Decode gains $\boldsymbol{\theta} = \boldsymbol{\ell} + (\bar{\boldsymbol{u}} - \boldsymbol{\ell}) \circ \boldsymbol{z}_k$; run all training closed-loop simulations and obtain the worst violation $\nu_{\max} = \max_s(\nu_{\mathrm{obs},s} + \nu_{\mathrm{uav},s})$;
    If $\nu_{\max} \le 10^{-10}$ and all per-scene cost terms are finite, set $J = \frac{1}{S}\sum_{s=1}^{S} \mathcal{J}_s$ and $F = J/(1+J)$; otherwise set $F = 1 + \nu_{\max}$;
  end
  If the generation is complete (Λ candidates evaluated), update the historical best candidate (lowest F, feasible or not), select the μ candidates with the lowest F, update the mean $\boldsymbol{m}$ by rank-weighted recombination, update the evolution paths $\boldsymbol{p}_\sigma, \boldsymbol{p}_c$, the covariance $\boldsymbol{C}$, and the step size σ; a truncated final generation does not update the distribution;
  if $\sigma < 10^{-6}$, 30 *consecutive generations without a new best candidate,* $\mathrm{cond}(\boldsymbol{C}) > 10^{12}$, *or* $\boldsymbol{C}$ *is non-finite* **then**
    Execute a restart: double Λ (up to 64), reset $\boldsymbol{C} = \boldsymbol{I}$, $\boldsymbol{p}_\sigma = \boldsymbol{p}_c = \boldsymbol{0}$, and $\sigma = 0.3$, and re-center $\boldsymbol{m}$ at the historical best candidate;
  end
end
return the feasible candidate with minimum training fitness F;

2.4.3 Tuning Protocol and Certificate Consistency

Each candidate is evaluated on six missions: two training instances from each of {dense, corridor, dynamic-target}. The budget is 816 candidate evaluations and validation seeds are never used for selection. The full-dimensional champion is selected by the pre-registered lowest-training-fitness rule among the feasible runs configured with training blocks 1–3 (runs 1–3 of the logged runs); the selected run is run 3. The resulting gains are listed in Table 3. Substitution in Eq. (51) gives $\Delta t_{\max} \approx 0.054\,\mathrm{s}$, above the implemented $\Delta t = 0.05\,\mathrm{s}$, so the discrete validation remains covered by the continuous-time UUB certificate.

3 Experimental Simulation

This section reports four experiments under a common two-scenario protocol, with the dense dynamic scene as the stress case. The fundamental-effectiveness study verifies the closed-loop operation of the tuned controller; the CMA-ES validation study compares offline tuning with default and degraded gains; the controller-baseline comparison evaluates the proposed architecture against three reduced controllers under an

Parameter	Default	Degraded	CMA-ES tuned
K_A	1.0000	1.0000	1.4878
K_R	1.4000	0.5000	1.4474
K_I	1.0000	1.0000	2.3012
K_s	2.0000	2.0000	1.2387
c	1.0000	1.0000	1.2685
η_o	0.0800	0.0300	0.1301
η_u	0.0500	0.0500	0.1951
K_G	0.6000	0.6000	1.1328
λ	1.0000	1.0000	2.9809
γ	0.0500	0.0500	0.1060

Table 3 Nominal, degraded, and CMA-ES tuned control gains.

identical tuning protocol; and the ablation study removes individual functional modules of the control law. Tuning details and optimized gains are given in Section 2.4; all validation uses seeds never seen during fitness evaluation.

3.1 Experimental Setup and Evaluation Metrics

To ensure reproducibility and enable fair comparison, this subsection fixes three sets of quantities shared by all experiments.

Table 4 lists the simulation environment parameters; the four safety/influence distances therein are shared by the controllers and the CMA-ES objective.

Parameter	Value	Physical Meaning	Parameter	Value	Physical Meaning
N	4	Number of follower UAVs	R_{comm}	3.2 m	Communication range
T	30 s	Total simulation duration	$d_{\text{obs}}^{\text{safe}}$	0.22 m	Obstacle safety distance
Δt	0.05 s	Simulation time step	$d_{\text{obs}}^{\text{inf}}$	1.45 m	Obstacle influence distance
$U_{\max}$	6.0 m/s^2	Maximum control input	$d_{\text{uav}}^{\text{safe}}$	0.65 m	Inter-UAV safety distance
$v_{\max}$	1.35 m/s	Maximum 3D speed before vertical clipping	$d_{\text{uav}}^{\text{inf}}$	1.25 m	Inter-UAV collision influence distance
$v_{z,\max}$	0.45 m/s	Maximum vertical velocity	Disturbance level	aggressive	Heterogeneous disturbance and impact disturbance
$[z_{\min}, z_{\max}]$	$[-1.40, 1.90]$ m	Altitude-control corridor	t_{pred}	0.5 s	Obstacle prediction horizon

Table 4 Basic simulation parameters of the experiments with $M = T/\Delta t + 1 = 601$ sampled states over 600 time intervals.

Formation offsets $\Delta \boldsymbol{d}_1 = [-1.5, 1.05, 0.30]^\top$, $\Delta \boldsymbol{d}_2 = [-1.5, -1.05, 0.30]^\top$, $\Delta \boldsymbol{d}_3 = [-3.0, 0.70, -0.30]^\top$, $\Delta \boldsymbol{d}_4 = [-3.0, -0.70, -0.30]^\top$ (m); initial follower–follower separations range from 1.40 to 2.38 m.

The Default, Degraded, and CMA-ES tuned values of the ten-dimensional gain vector $\boldsymbol{\theta}$ in Eq. (55) are consolidated in Table 3; the fixed prediction horizon is included

in Table 4. Unless stated otherwise, all experiments below use the CMA-ES tuned gains.

Table 5 defines the unified evaluation metrics; all quantitative comparisons are computed over NM samples with the safety thresholds from Table 4. The two success indicators are averaged over the random seeds of each experiment to yield the reported success rates.

Metric	Symbol	Definition
Minimum obstacle distance	$d_{\min}^{\text{obs}}$	$\min_{k,\, i\in\{L,1,\ldots,N\}} d_{\text{obs},i}(k)$
Minimum inter-UAV distance	$d_{\min}^{\text{uav}}$	$\min_{k,\, i\neq j} d_{ij}(k)$
Tracking RMSE	RMSE	$\sqrt{\frac{1}{NM}\sum_{k=1}^{M}\sum_{i=1}^{N}\|\boldsymbol{p}_i(k)-\boldsymbol{p}_{d,i}(k)\|^2}$
Control energy	E_u	$\sum_{k,i}\|\boldsymbol{u}_i(k)\|^2\Delta t$
Control smoothness	$J_{\Delta u}$	$\sum_{k=2}^{M}\sum_{i=1}^{N}\|\boldsymbol{u}_i(k)-\boldsymbol{u}_i(k-1)\|^2$
Input saturation rate	r_{sat}	fraction of NM samples with $\|\boldsymbol{u}_i(k)\| > 0.95\,U_{\max}$
Obstacle-avoidance success	S_{obs}	$\mathbb{I}[d_{\min}^{\text{obs}} \geq d_{\text{obs}}^{\text{safe}}]$
Joint success rate	S_{joint}	$\mathbb{I}[d_{\min}^{\text{obs}} \geq d_{\text{obs}}^{\text{safe}} \wedge d_{\min}^{\text{uav}} \geq d_{\text{uav}}^{\text{safe}}]$

Table 5 Unified evaluation metrics of the experiments.

Unless stated otherwise, every experiment is evaluated over $N_s = 30$ independent seeds (1001–1030). A common horizontal obstacle-field jitter $\boldsymbol{j}_{0,q} \sim \mathcal{N}(\boldsymbol{0}, 0.15^2\boldsymbol{I}_2)$ is applied first; each seed then adds an independent $\boldsymbol{j}_{r,q} \sim \mathcal{N}(\boldsymbol{0}, 0.15^2\boldsymbol{I}_2)$ to every obstacle center. Thus the seed-conditioned perturbation relative to the common base field has standard deviation 0.15 m per horizontal axis. Initial follower positions receive $\boldsymbol{p}_i(0) = \boldsymbol{p}_L(0) + \Delta\boldsymbol{d}_i + 0.24\,\boldsymbol{\xi}_i$, with $\boldsymbol{\xi}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_3)$, corresponding to perturbation scale 1.2×0.20. The aggressive disturbance profile is activated and the discrete white-noise acceleration is $0.03\,\boldsymbol{w}_i[k]/\sqrt{\Delta t}$, $\boldsymbol{w}_i[k] \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_3)$. These validation seeds are disjoint from the training instances used during fitness evaluation. Paired statistical tests always compare configurations on identical seeds.

3.2 Fundamental Effectiveness in the Dense Dynamic Scenario

Using the run-3 CMA-ES gains fixed in Section 2.4, this subsection first verifies the fundamental closed-loop operation of the complete controller in the dense dynamic scene. This experiment verifies the fundamental closed-loop operation of the proposed controller in the dense dynamic obstacle environment, the most demanding scene of this work. The simulation system consists of 1 leader UAV and 4 follower UAVs. The followers adopt a layered rhombus expected formation and operate in 3D space for 30 s. The mission divides into five phases: 0–3 s take-off, 3–7 s cruise climbing, 7–18 s main obstacle avoidance, 18–24 s descent, and 24–30 s landing.

3.2.1 Obstacle Scene Configuration

The dense dynamic scene contains 15 obstacles—8 spheres, 3 boxes, 2 cylinders, 1 ellipsoid, and 1 torus—9 of which move periodically; a lighter standard scenario is used in the CMA-ES validation and controller-baseline comparison. Fig. 4 shows the complete 3D trajectories in the dense dynamic scene for a representative seed: the leader and four followers complete the full mission, with follower trajectories deviating locally from the desired formation during obstacle avoidance and then rejoining the

leader trajectory. Fig. 5 shows the corresponding top and side views: in the top view the followers weave around the obstacle field, and in the side view the trajectories remain inside the altitude corridor on this seed.

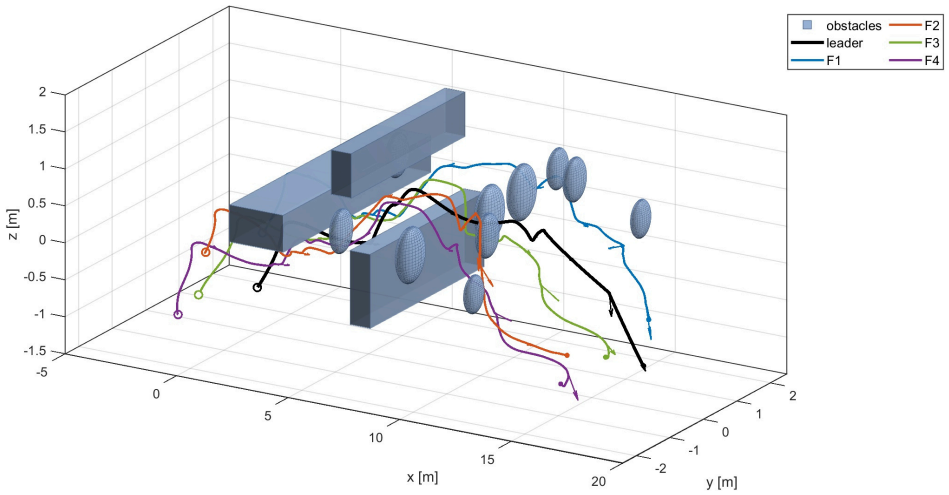
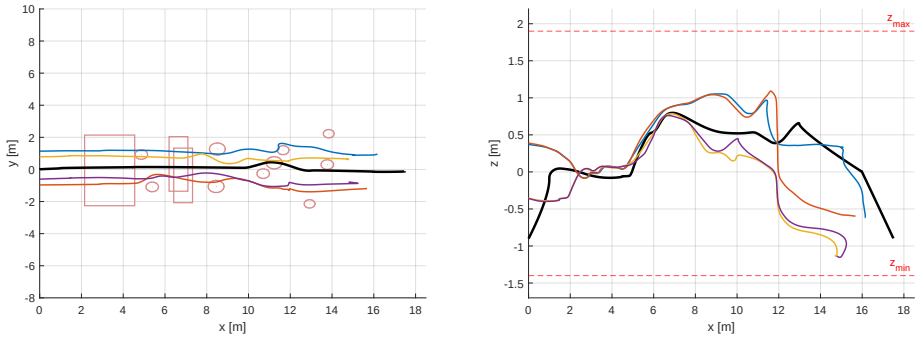


Fig. 4 3D trajectories of the UAV swarm in the dense dynamic scene.



(a) Top view (dotted: dynamic obstacle sweep) (b) Side view (altitude corridor)

Fig. 5 Top and side views of the trajectories in Fig. 4.

3.2.2 Simulation Results

Table 6 lists the quantitative results over the 30 seeds. The obstacle-avoidance success rate reaches 0.900: 27 of 30 seeds keep the mission-level clearance above $d_{\text{obs}}^{\text{safe}} = 0.22\,\text{m}$ with a mean margin of 0.029 m, and even the worst seed violates the threshold by only 1.2 cm ($d_{\text{min}}^{\text{obs}} = 0.208\,\text{m}$). The mean minimum inter-UAV distance is $0.722 \pm 0.150\,\text{m}$, above the 0.65 m threshold on average, yielding a joint success rate of 0.667. Control inputs remain essentially unsaturated: the saturation rate is 0.0011, i.e., 0.11% of all samples, and the largest applied norm is reported as $6.00\,\text{m/s}^2$ at the stated reporting precision.

Table 6 Quantitative results in the dense dynamic scene over 30 seeds.

Category	Metric	Value
Safety	$d_{\min}^{\text{obs}}$/m	0.249 ± 0.020
	Margin/m	0.029 ± 0.020
	$d_{\min}^{\text{uav}}$/m	0.722 ± 0.150
	S_{obs}	0.900
	S_{joint}	0.667
Tracking and control	RMSE/m	0.879 ± 0.290
	$J_{\Delta u}$	337.1 ± 151.5
	E_u	160.2 ± 69.8

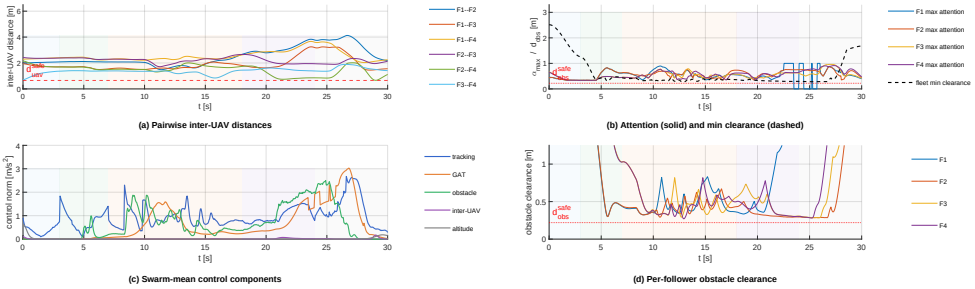


Fig. 6 Time histories in the dense dynamic scene: (a) the six pairwise follower separations; (b) each follower's maximum attention weight and the fleet minimum obstacle clearance; (c) the five swarm-mean control-component norms; and (d) the four per-follower obstacle clearances. Dashed and dotted red lines denote the corresponding safety thresholds.

Fig. 6 summarizes the key per-mission time histories.

(a) Inter-UAV distances. The six pairwise follower distances start at their smallest values around take-off (0.663 m, already above $d_{\text{uav}}^{\text{safe}}$), expand up to 4.1 m while the swarm deforms around obstacles, and never cross the threshold during the mission of this seed.

(b) Attention weights and obstacle clearance. The maximum attention weight $\alpha_{\max}$ starts near 0.47, varies over the full $[0, 1]$ range, and exhibits isolation events in which attention concentrates on a single neighbor ($\alpha_{\max} = 1.0$ near $t \approx 22.6$ s). Meanwhile, the global minimum obstacle distance (dashed) stays above 0.22 m throughout, with its minimum of 0.271 m occurring at $t \approx 13.0$ s in the main avoidance phase.

(c) Control components. Only the five swarm-mean component norms are plotted; the total input is not superimposed to avoid obscuring these components. The GAT-compensation term is the largest contributor at its peak ($3.03\,\text{m/s}^2$ near $t \approx 26.8$ s), followed by the tracking term ($2.67\,\text{m/s}^2$) and the obstacle-avoidance term ($2.50\,\text{m/s}^2$), all peaking in the final recovery–landing segment when the swarm restores the formation while re-crossing the obstacle field. The collision-avoidance term stays below $0.17\,\text{m/s}^2$, and the altitude term acts mainly during take-off. The total per-follower input is likewise omitted from the per-follower decomposition below; its representative-seed peak is $5.80\,\text{m/s}^2$ near $t \approx 26.6$ s.

(d) Per-follower clearance. All four per-follower clearances stay above the threshold, with per-follower minima between 0.271 and 0.300 m.

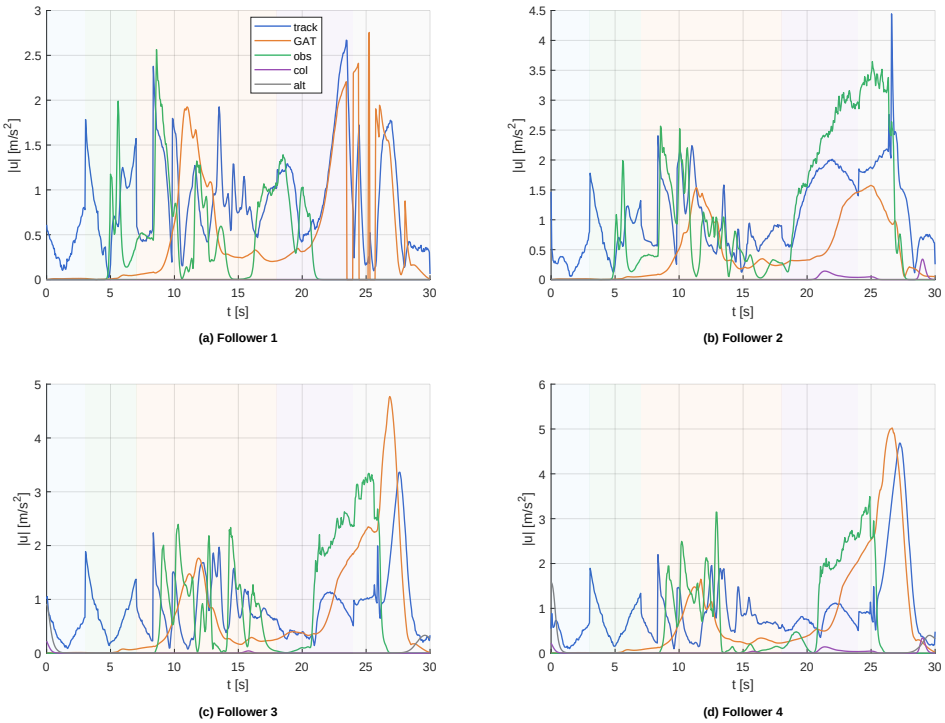


Fig. 7 Control-component decomposition of the four follower UAVs over the mission.

Fig. 7 decomposes the control inputs of the four followers into the five component norms; the total command is omitted to keep the component behavior visible. The tracking, GAT-compensation, and obstacle-avoidance components dominate and return to small values after formation recovery, whereas the inter-UAV collision-avoidance and altitude terms remain small throughout, consistent with the large separation margins and the absence of altitude violations.

3.3 Validation of CMA-ES Offline Tuning

The tuned gains were evaluated in the standard and dense scenarios over the 30 independent seeds, with the untuned default gains and a degraded gain set as references; the degraded gains (Table 3) emulate gain miscalibration after maintenance or parameter drift. Table 7 reports mission-level safety, tracking, and smoothness statistics.

Two observations follow. First, the degraded gains collapse obstacle safety to zero in *both* scenarios—every seed violates the obstacle threshold—even though their energy and smoothness costs are the lowest; safe gains thus occupy a small region of the ten-dimensional space, which motivates principled offline tuning. Second, in the dense scene, tuning raises the joint success rate from 0.033 to 0.667 (20×): the mean minimum inter-UAV distance increases from 0.539 m to 0.722 m, crossing the 0.65 m threshold, while RMSE decreases from 1.298 m to 0.879 m and obstacle clearance improves from 0.230 m to 0.249 m. The smoothness cost rises moderately (+13%), and the saturation rate remains 0.0011. In the standard scene the default gains are already

Table 7 Validation of degraded, default, and CMA-ES tuned gains over 30 seeds.

Scenario	Configuration	$d_{\min}^{\text{obs}}$/m	$d_{\min}^{\text{uav}}$/m	RMSE/m	S_{obs}	S_{joint}	$J_{\Delta u}$
Standard	Degraded	0.163	0.637	0.9577	0.000	0.000	225
Standard	Default	0.243	0.535	1.0981	1.000	0.167	281
Standard	CMA-ES	0.250	0.569	1.0454	1.000	0.333	349
Dense	Degraded	0.155	0.614	0.8970	0.000	0.000	289
Dense	Default	0.230	0.539	1.2979	0.933	0.033	299
Dense	CMA-ES	0.249	0.722	0.8791	0.900	0.667	337

nearly sufficient and tuning mainly improves tracking (RMSE 1.098 $\rightarrow$ 1.045 m) while preserving $S_{\text{obs}} = 1.0$; the inter-UAV margin (0.569 m), however, remains below the threshold, a scene-level bottleneck that reappears in the controller-baseline comparison and ablation study. Thus offline tuning improves the safety–performance balance without changing the online control law or its stability certificate.

3.4 Controller Baseline Comparison

This experiment isolates the contribution of the controller *architecture* from gain tuning. Three reduced controllers are derived from the proposed control law by removing channels; all four controllers are tuned by CMA-ES with the same 816-evaluation budget, the same three-block training instance pool, and the same feasibility-first champion-selection rule, and each controller tunes only the gains of its active channels (3, 7, 7, and 10 dimensions, respectively). The population size follows the CMA-ES dimension default, giving $\Lambda = 7, 9, 9, 10$ for B1–Ours. Validation uses the same 30 seeds as the previous experiments in both scenarios. Table 8 defines the baselines.

Table 8 Controller baseline configurations.

Module	B1	B2	B3	Ours
Backstepping tracking $\boldsymbol{u}_{\text{track}}$	✓	✓	✓	✓
Obstacle APF $\boldsymbol{u}_{\text{obs}}$	–	✓	✓	✓
Inter-UAV APF $\boldsymbol{u}_{\text{col}}$	–	✓	✓	✓
Conflict projection ACP (effective)	Inactive	Active	Active	Active
Smooth saturation	✓	✓	✓	✓
Predictive avoidance	–	–	✓	✓
Adaptive gain	–	–	✓	✓
GAT attention compensation	–	–	–	✓
Safety gate	–	–	–	✓
Tuned dimensions	3	7	7	10
Feasible champion found	No	No	Yes	Yes

B1 is the backstepping tracking controller; B2 augments it with obstacle and inter-UAV APF terms; B3 further adds the adaptive-gain and predictive-avoidance mechanisms and differs from the complete architecture only in the GAT compensation channel and its safety gate. ACP and smooth saturation are common implementation post-processors for the APF-equipped controllers. The tuning outcome itself is informative: within the identical budget, B1 and B2 found no feasible champion in any of

their three tuning blocks, whereas B3 and the complete architecture did; for these two baselines the lowest-training-fitness (infeasible) candidates are retained for validation, which is the most favorable tuning outcome their architectures admit. A missing avoidance or compensation channel therefore cannot be rescued by gain tuning alone—the bottleneck is architectural, not parametric. Table 9 reports the validation results.

Table 9 Controller comparison over 30 seeds in both scenarios.

Scenario	Method	$d_{\min}^{\text{obs}}$/m	$d_{\min}^{\text{uav}}$/m	RMSE/m	S_{obs}	S_{joint}	$J_{\Delta u}$	E_u	r_{sat}
Standard	B1 Backstepping only	−0.273	1.067	0.4047	0.000	0.000	84	31.8	0.0000
Standard	B2 Backstepping+APF	0.185	0.649	1.3794	0.067	0.067	2080	805.4	0.0779
Standard	B3 No GAT/gate	0.263	0.602	1.3302	1.000	0.433	187	79.6	0.0000
Standard	Ours (full)	0.250	0.569	1.0454	1.000	0.333	349	170.5	0.0014
Dense	B1 Backstepping only	−0.275	1.066	0.4516	0.000	0.000	417	208.8	0.0048
Dense	B2 Backstepping+APF	0.155	0.594	1.8597	0.000	0.000	3771	1183.9	0.1107
Dense	B3 No GAT/gate	0.232	0.649	2.4027	0.867	0.333	496	179.5	0.0015
Dense	Ours (full)	0.249	0.722	0.8791	0.900	0.667	337	160.2	0.0011

Dense scene. B1 traverses the obstacle region without an avoidance channel: its mean obstacle clearance is −0.275 m and its obstacle-avoidance success rate is zero. B2 adds APF repulsion but still gives a mean clearance of only 0.155 m, a mean inter-UAV distance of 0.594 m, an 11.1% saturation rate, and the largest smoothness and energy costs among the four controllers. B3 is the strongest reference: predictive avoidance raises its obstacle-avoidance success rate to 0.867, but its RMSE is 2.403 m and its mean inter-UAV distance is only 0.649 m. The complete controller keeps a mean clearance of 0.249 m, increases the mean inter-UAV distance to 0.722 m, and reduces RMSE to 0.879 m. Its joint success rate doubles from 0.333 for B3 to 0.667, while the smoothness cost falls from 496 to 337 and the energy cost falls from 179.5 to 160.2. The dense scene repeatedly forces the swarm to deform in order to pass obstacles, so formation recovery and avoidance must act simultaneously. Because B3 lacks the bounded GAT compensation channel and the safety gate, its followers remain loosely distributed after avoidance. The complete architecture restores the formation through GAT compensation when the margin is large and suppresses that compensation near obstacles through the safety gate, which explains its clearer advantage in the dense scene.

Standard scene. B3 and the complete controller both achieve an obstacle-avoidance success rate of one in the standard scene. B3 gives a mean obstacle clearance of 0.263 m and a joint success rate of 0.433, whereas the complete controller gives 0.250 m and 0.333. The architectural advantage in this sparser scene is therefore concentrated in tracking rather than safety: RMSE decreases from 1.330 m for B3 to 1.045 m for the complete controller. The standard scene leaves more room around the obstacles, so the classical avoidance channels can maintain safety with less formation deformation. The benefit of GAT compensation is consequently smaller than in the dense scene, while the inter-UAV margin remains the binding constraint for both controllers.

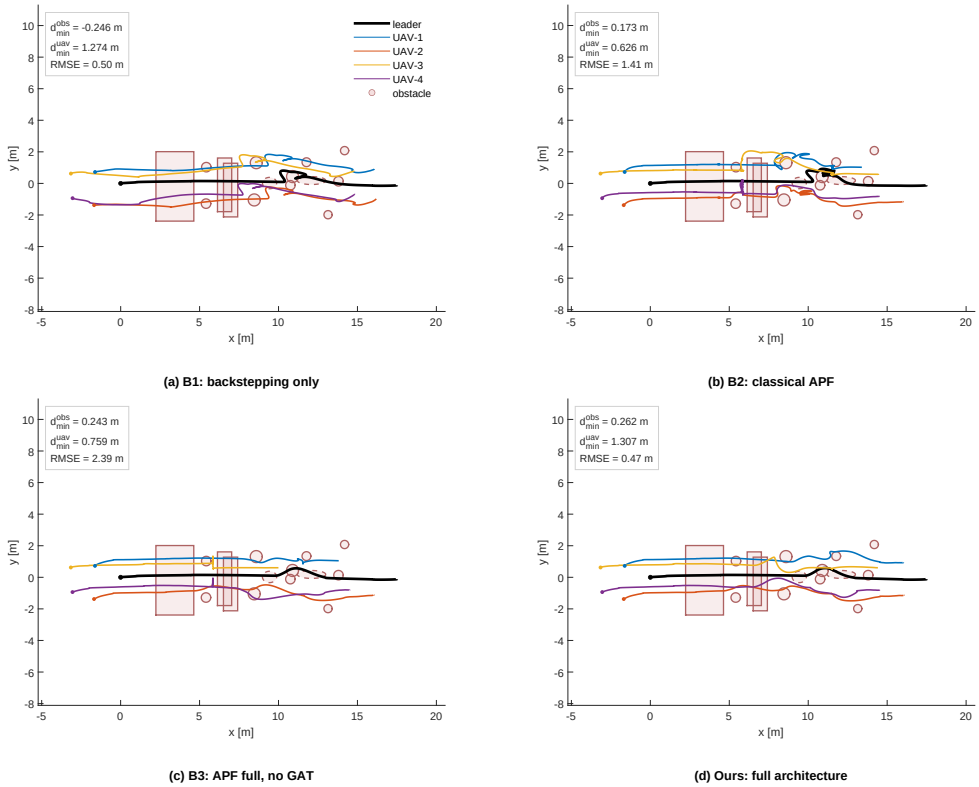


Fig. 8 Top-view trajectories of the four controllers.

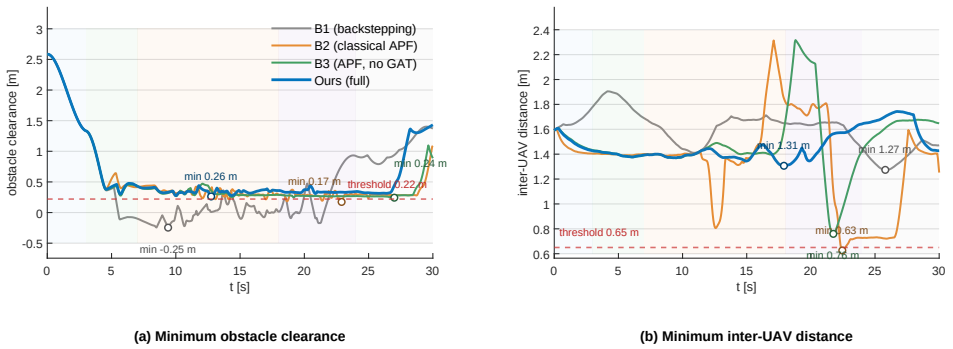


Fig. 9 Minimum obstacle clearance and inter-UAV distance of the four controllers.

Figs. 8 and 9 visualize a representative dense-scene seed. B1 reaches a clearance of $-0.246\,\mathrm{m}$ near $t = 9.4\,\mathrm{s}$ and therefore passes through the obstacle field. B2 avoids geometric penetration but reaches only $0.173\,\mathrm{m}$ clearance and $0.626\,\mathrm{m}$ minimum separation, while its time histories show strong oscillation. B3 is safe on this seed, with $0.243\,\mathrm{m}$ clearance and $0.759\,\mathrm{m}$ minimum separation, but its RMSE is $2.39\,\mathrm{m}$ and its followers remain loosely distributed. The complete controller reaches $0.262\,\mathrm{m}$ clearance, $1.31\,\mathrm{m}$ minimum separation, and $0.47\,\mathrm{m}$ RMSE. Its clearance and separation curves

stay farther from the thresholds and recover more smoothly after the avoidance phase. With the tuner, budget, and seeds held constant, the bounded attention compensation with its safety gate is therefore the component that converts an avoidance-capable architecture into one that also maintains tight formation tracking.

3.5 Ablation Study

To quantify the contribution of each functional module, nine single-module ablation variants are constructed from the complete controller at the tuned operating point: No-Track, No-GAT, No-Attention, No-Gate, No-Obs, No-Predict, No-AdaptGain, No-Col, and No-Altitude; each variant disables exactly one module. All variants are evaluated on the same 30 dense-scene seeds, and per-seed paired differences against the Full configuration are tested with Wilcoxon signed-rank tests under Benjamini–Hochberg false-discovery-rate correction (reported as q-values). Table 10 and Figs. 10–11 summarize the results.

Table 10 Ablation results in the dense scene over 30 seeds.

Config	$d_{\min}^{\text{obs}}$/m	$d_{\min}^{\text{uav}}$/m	RMSE/m	$J_{\Delta u}$	E_u	S_{obs}	S_{joint}
Full	0.249 ± 0.020	0.722 ± 0.150	0.879 ± 0.290	337.1	160.2	0.900	0.667
No-Obs	-0.261 ± 0.035	1.227 ± 0.175	0.082 ± 0.010	123.3	44.8	0.000	0.000
No-Predict	0.124 ± 0.043	0.512 ± 0.197	0.881 ± 0.230	1534.6	648.0	0.000	0.000
No-Track	0.224 ± 0.035	0.655 ± 0.276	7.329 ± 2.777	651.0	187.4	0.533	0.233
No-GAT	0.243 ± 0.019	0.704 ± 0.221	2.423 ± 0.594	661.6	211.4	0.933	0.533
No-Attention	0.246 ± 0.022	0.685 ± 0.110	0.856 ± 0.241	316.3	119.3	0.867	0.567
No-Gate	0.244 ± 0.018	0.713 ± 0.146	0.843 ± 0.269	363.2	160.4	0.900	0.667
No-AdaptGain	0.232 ± 0.018	0.693 ± 0.101	0.894 ± 0.224	395.5	166.7	0.800	0.633
No-Col	0.249 ± 0.020	0.686 ± 0.171	0.891 ± 0.295	360.8	163.7	0.900	0.467
No-Altitude	0.249 ± 0.020	0.728 ± 0.148	0.879 ± 0.290	336.3	159.3	0.900	0.667

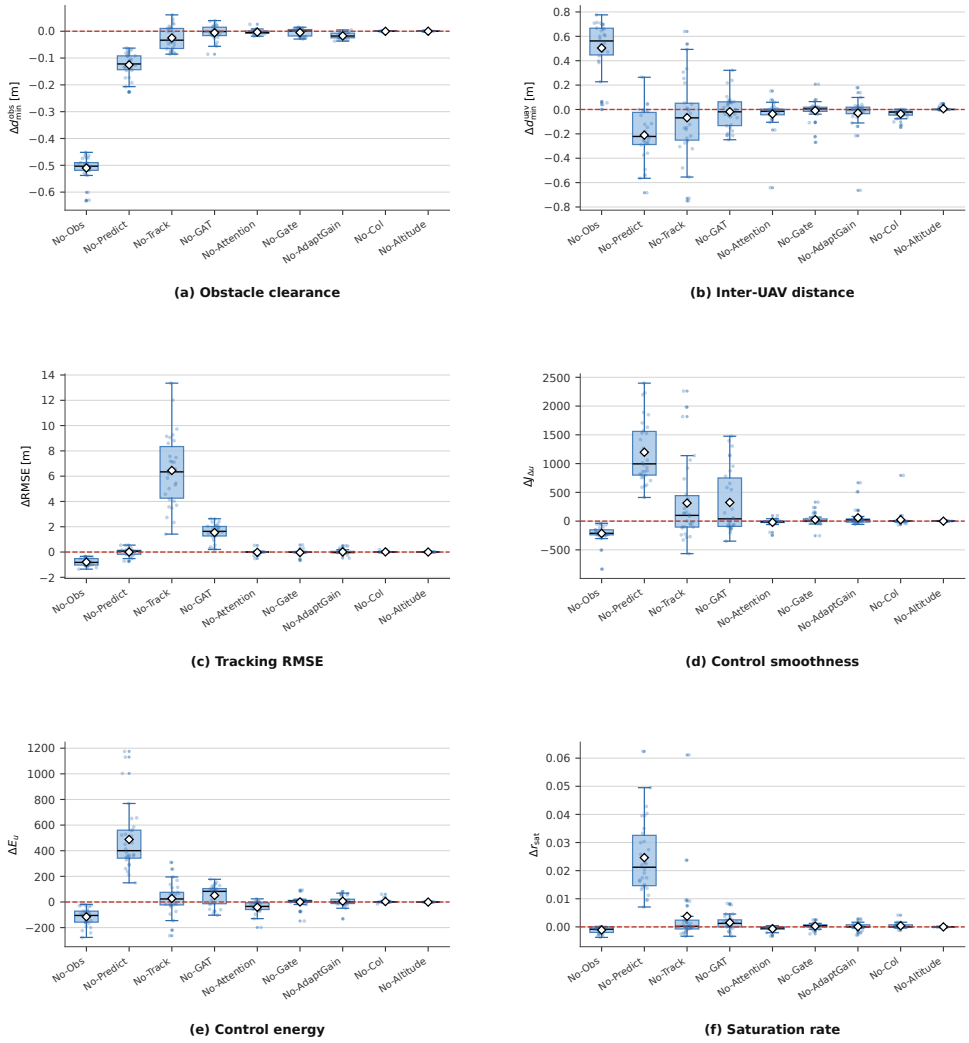


Fig. 10 Paired differences of the nine ablation variants in the dense scene.

Safety-critical modules. The obstacle-avoidance and predictive layers dominate the safety panels of Fig. 10. Disabling obstacle avoidance lowers the paired obstacle-clearance difference by about 0.51 m and shifts the entire distribution below zero, showing that the followers enter the obstacle region rather than merely losing margin. Its apparently favorable RMSE difference is therefore misleading: the followers reduce formation error by cutting through the obstacle field instead of bypassing it. Disabling predictive avoidance lowers obstacle clearance by about 0.13 m and inter-UAV separation by about 0.21 m, while increasing the smoothness-cost difference by more than 1190 units and the energy-cost difference by about 488 units. This pattern indicates late reactions to moving obstacles and consequently sharper, more expensive avoidance maneuvers.

Tracking and GAT compensation. The tracking and compensation channels mainly affect the performance panels. Disabling nominal tracking produces the largest RMSE degradation in Fig. 10, with a paired increase of about 6.45 m, and also reduces both obstacle clearance and inter-UAV separation. Thus, stable formation keeping is not only a tracking objective but also an indirect mechanism for maintaining safe follower separation. Disabling GAT produces a smaller safety change but increases RMSE by about 1.54 m and raises the smoothness-cost difference by about 325 units. This degradation matches the behavior of the strongest controller baseline without GAT: after avoidance-induced deformation, the swarm recovers its formation more slowly and with larger control effort. The GAT channel therefore restores tracking performance during avoidance without being the primary source of obstacle safety.

Gate and attention aggregation. Removing the safety gate has little effect on the paired safety differences and slightly reduces RMSE, but it increases the smoothness-cost difference by about 26 units. In nominal operation, therefore, the gate should be interpreted as a bounded-authority safeguard rather than a mechanism that improves tracking performance. Replacing attention aggregation by uniform aggregation produces only small differences in the safety and RMSE panels. Hence, the observed benefit comes mainly from the bounded compensation channel and less from the specific softmax weighting of neighbors.

Secondary modules and the inter-UAV bottleneck. Removing the adaptive gain lowers the obstacle-clearance difference by about 0.017 m and increases both smoothness and energy costs, showing that distance-based authority redistribution provides useful margin and damping. Removing inter-UAV collision avoidance produces the clearest negative effect in the inter-UAV panel, about 0.036 m, while the obstacle-clearance and RMSE differences remain small. Thus, inter-UAV separation is the tightest mission constraint: it is maintained jointly by nominal tracking, collision repulsion, and GAT compensation, but with limited surplus. The altitude variant lies close to zero in all six panels because the optimized trajectories already remain inside the altitude corridor; its role is constraint enforcement rather than performance improvement in this scene.

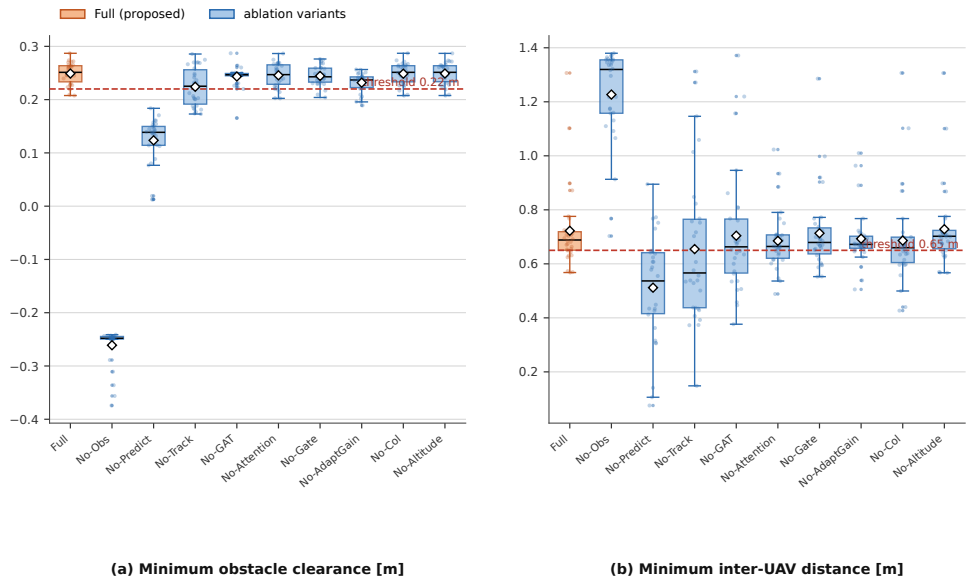


Fig. 11 Distributions of the two safety metrics for the Full configuration and nine ablation variants.

Fig. 11 shows the 30-seed distributions. The No-Obs box lies entirely below zero clearance and No-Predict entirely below the 0.22 m threshold, while No-Track spans the widest inter-UAV spread; the inter-UAV boxes of No-Track, No-GAT, No-Attention, and No-Col all reach below the 0.65 m threshold. Together with the controller-baseline comparison, the ablation identifies obstacle avoidance, predictive avoidance, nominal tracking, and bounded GAT compensation as the safety- and performance-critical modules, and locates the remaining joint-success losses in the inter-UAV separation margin.

4 Conclusion

This paper presented a safety-gated graph-attention compensation backstepping framework for multi-UAV formation flight in obstacle-rich environments. A risk-weighted graph-attention module acts as an additive compensator; an obstacle-distance authority factor jointly reduces tracking gain and increases obstacle repulsion; a safety-conflict projection removes tracking components that directly oppose near-obstacle repulsion; and a safety gate redistributes authority among tracking, obstacle avoidance, inter-UAV anti-collision, altitude constraint, and compensation channels. Smooth saturation enforces the acceleration bound. The Lyapunov analysis of the smooth nominal law certifies uniform ultimate boundedness, whereas the conflict projection and saturation are implementation-level safety post-processors evaluated in simulation. The ten coupled gains were tuned offline with restart-enabled CMA-ES using a fixed budget, a feasibility-first objective, and validation-independent selection; the selected gains satisfy the discretization condition associated with the stability certificate.

In a 30-seed dense dynamic scene, the tuned controller achieved a mean minimum obstacle clearance of 0.249 ± 0.020 m, obstacle-avoidance success of 90.0%, joint success of 66.7%, and RMSE of 0.879 m. CMA-ES tuning increased dense-scene joint success from 3.3% to 66.7%. Controller comparisons showed that pure backstepping cannot maintain obstacle safety, the APF baseline suffers infeasible or saturated commands, and removing GAT compensation increases RMSE from 0.88 m to 2.40 m. Ablation experiments identify obstacle avoidance and predictive avoidance as safety-critical modules, and nominal tracking and bounded GAT compensation as performance-critical modules.

The inter-UAV spacing constraint remains an architectural bottleneck in the densest scenes, safety improvement trades against smoothness and energy, the certificate does not cover intervals in which projection or saturation is active, and validation is numerical with reliable delay-free leader broadcast. Future work will address dedicated inter-UAV safety filters, multi-objective tuning, communication delays, packet loss, switching topologies, and hardware-in-the-loop or outdoor flight validation.

Author contributions. Hao Wang: conceptualization, methodology, and original draft preparation. Guhao Zhao: supervision, review and editing, and correspondence. Yulong Cao: writing, validation, and figure preparation. All authors reviewed and approved the final manuscript.

Code availability. The custom MATLAB simulation code used in this study is available from the corresponding author on reasonable request.

Data availability. The numerical values supporting the tables and figures are included in this article. The underlying per-seed simulation data are available from the corresponding author on reasonable request.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests. The authors declare no competing interests.

References

[1] Dentler, J., Kannan, S., Bezzaoucha, S., Olivares-Mendez, M. A. & Voos, H. Model predictive cooperative localization control of multiple UAVs using potential function sensor constraints. *Autonomous Robots* **43**, 153–178 (2019).

[2] Wang, T., Zhao, S., Xia, Y., Pan, Z. & Tian, H. Consensus control of large-scale UAV swarm based on multi-layer graph. *Drones* **6**, 402 (2022).

[3] Jiang, Y., Bai, T. & Wang, Y. Formation control algorithm of multi-UAVs based on alliance. *Drones* **6**, 431 (2022).

[4] Dai, S. *et al. Real-time collision-free planning and tracking control for multiple UAVs based on distributed model predictive control.* Proceedings of the 2023 China Automation Congress, 4348–4355 (IEEE, 2023).

[5] Niu, Z., Jia, X. & Yao, W. Communication-free MPC-based neighbors trajectory prediction for distributed multi-UAV motion planning. *IEEE Access* **10**, 13481–13489 (2022).

[6] Huo, D., Dai, L., Chai, R., Xue, R. & Xia, Y. Collision-free model predictive trajectory tracking control for UAVs in obstacle environment. *IEEE Transactions on Aerospace and Electronic Systems* **59**, 2920–2932 (2023).

[7] Niu, K. *et al. Agile obstacle avoidance control based on dynamic velocity obstacle perception.* Proceedings of the 2025 IEEE International Conference on Unmanned Systems, 484–491 (IEEE, 2025).

[8] Chen, Y.-W., Chiang, M.-L., Kuo, G.-R., Chung, C.-J. & Fu, L.-C. *Model predictive control with reference path planning for multi-UAV formation control system.* Proceedings of the 2024 American Control Conference, 279–284 (IEEE, 2024).

[9] Zhu, X., Liang, Y. & Yan, M. A flexible collision avoidance strategy for the formation of multiple unmanned aerial vehicles. *IEEE Access* **7**, 140743–140754 (2019).

[10] Zhang, P., He, Y., Wang, Z., Li, S. & Liang, Q. Research on multi-UAV obstacle avoidance with optimal consensus control and improved APF. *Drones* **8**, 248 (2024).

[11] Ma, B., Ji, Y. & Fang, L. A multi-UAV formation obstacle avoidance method combined with improved simulated annealing and an adaptive artificial potential field. *Drones* **9**, 390 (2025).

[12] Li, Y. *et al.* Multi-UAV obstacle avoidance and formation control in unknown environments. *Drones* **8**, 714 (2024).

[13] Tao, J., Guo, S. & Shang, Y. An effective strategy for collision avoidance of multiple UAVs with unknown acceleration. *IEEE Access* **11**, 112600–112619 (2023).

[14] Lv, X., Peng, C. & Ma, J. Control barrier function-based collision avoidance guidance strategy for multi-fixed-wing UAV pursuit-evasion environment. *Drones* **8**, 415 (2024).

[15] Fu, J., Lin, X. & Wen, G. Robust obstacle avoidance and safe formation tracking control for multiple fixed-wing UAVs based on high-order control barrier functions. *Robot* **47**, 85–98 (2025).

[16] Wei, Z., Zhang, X., Song, Y. & Guo, R. Formation control for UAVs considering safety constraints based on control barrier functions with switched trajectories and switching communication topologies. *Sensors* **26**, 1477 (2026).

[17] Liu, Z., Cao, Y., Chen, J. & Li, J. A hierarchical reinforcement learning algorithm based on attention mechanism for UAV autonomous navigation. *IEEE Transactions on Intelligent Transportation Systems* **24**, 13309–13320 (2023).

[18] Hamissi, A., Dhraief, A. & Sliman, L. *GADCA: A MARL-based graph-attentive decentralized collision avoidance framework for UAV in UTM.* Proceedings of the 2026 IEEE 23rd Consumer Communications and Networking Conference, 1–8 (IEEE, 2026).

[19] Yang, Q., Lu, J., Zhang, Y. & Shao, S. Multi-UAV cooperative path planning via graph neural network and proximal policy optimization. *IEEE Access* **13**, 193575–193588 (2025).

[20] Liu, Y. *et al.* Flexible multi-UAV formation control via integrating deep reinforcement learning and affine transformations. *Aerospace Science and Technology* **157**, 109812 (2025).

[21] Zhang, J., Yan, J. & Zhang, P. Multi-UAV formation control based on a novel back-stepping approach. *IEEE Transactions on Vehicular Technology* **69**, 2437–2447 (2020).

[22] Muslimov, T. Z. & Munasypov, R. A. Adaptive decentralized flocking control of multi-UAV circular formations based on vector fields and backstepping. *ISA Transactions* **107**, 143–159 (2020).

[23] Shi, Y., Wang, R., Tong, R. & Zeng, Y. *Tracking control and obstacle avoidance of quad-rotor UAV based on high-order fully actuated system approach.* 2024 3rd Conference on Fully Actuated System Theory and Applications (FASTA), 695–700 (IEEE, 2024).

[24] Li, J., Zhang, C., Li, B. & Zhang, J. *Trajectory tracking of quadrotor unmanned aerial vehicle based on adaptive backstepping sliding mode control.* 2024 36th Chinese Control and Decision Conference (CCDC), 3619–3624 (IEEE, 2024).

[25] Veličković, P. *et al. Graph attention networks.* Proceedings of the 6th International Conference on Learning Representations (2018).

[26] Hansen, N. & Ostermeier, A. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation* **9**, 159–195 (2001).

[27] Hansen, N. The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772* (2016).

[28] Auger, A. & Hansen, N. *A restart CMA evolution strategy with increasing population size.* Proceedings of the 2005 IEEE Congress on Evolutionary Computation, Vol. 2, 1769–1776 (2005).