Toward Uncertainty-Aware CBF Safety Filters

Hyo-Sang Shin

EuroGNC 2026
Pre-Conference Tutorial Workshop

Autonomous and Intelligent Systems Lab.
KAIST

4 May 2026

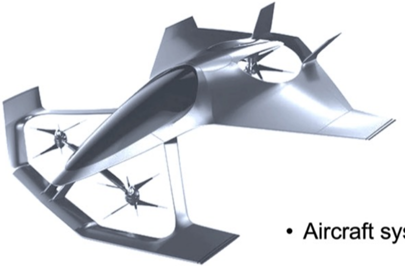
Outline

Outline

Research Background

- Area of expertise
  - Aerospace Guidance, Navigation, Control (GNC)
  - Backbone: Systems and Control
- Position
  - Professor at KAIST
  - Adjunct Professor of Guidance, Navigation, and Control at Cranfield University
  - Head of Autonomous and Intelligent Systems (AIS) Lab.

Research Themes

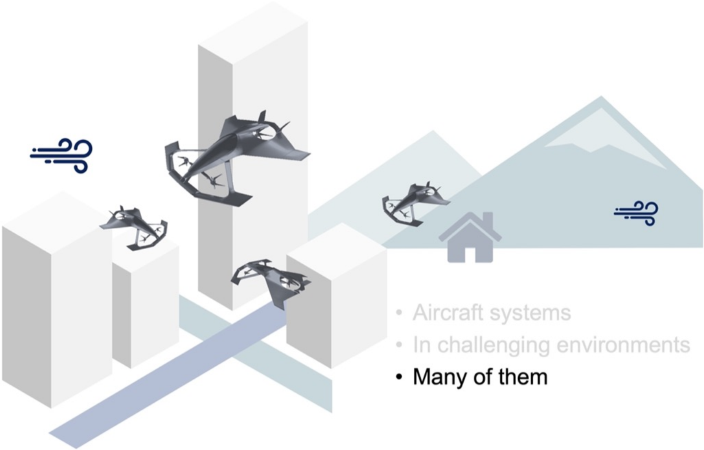


- Aircraft systems

Research Themes

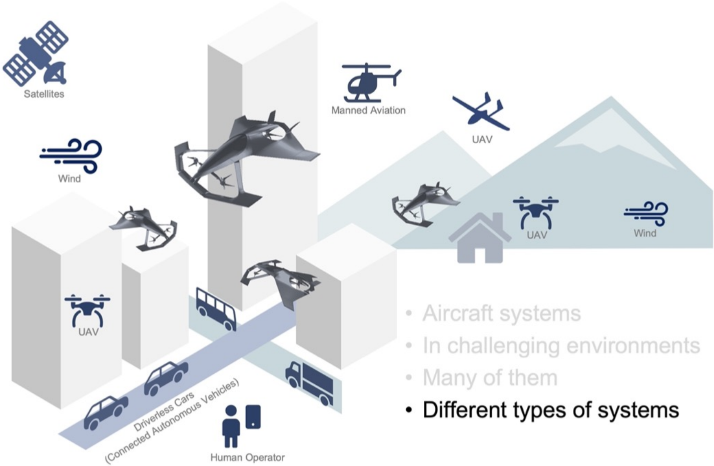
- Aircraft systems
- In challenging environments

Research Themes



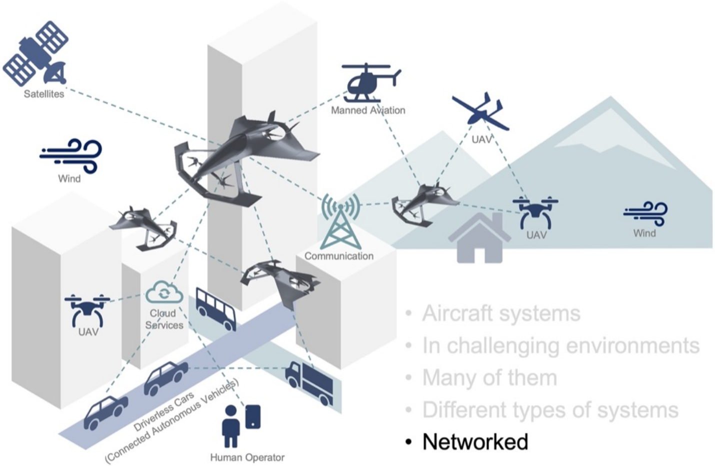
- Aircraft systems
- In challenging environments
- Many of them

Research Themes



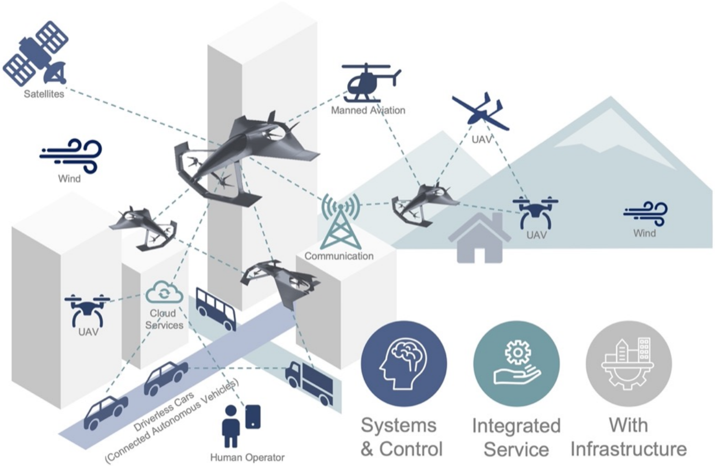
- Aircraft systems
- In challenging environments
- Many of them
- **Different types of systems**

Research Themes



- Aircraft systems
- In challenging environments
- Many of them
- Different types of systems
- Networked

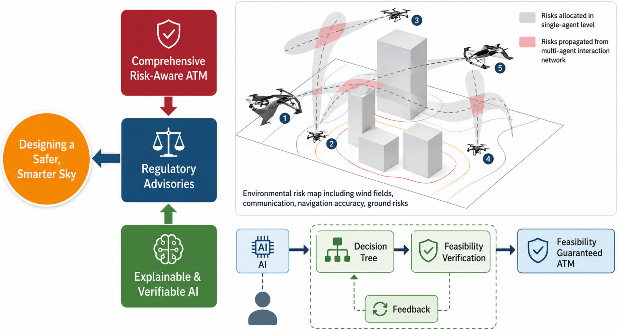
Research Themes



Outline

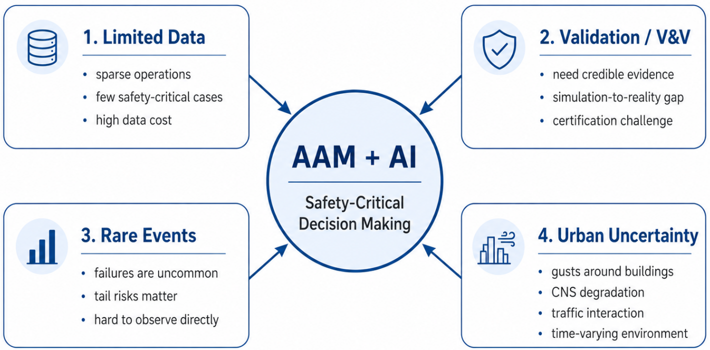
From AI to Assured Autonomy in Aerospace

- Aerospace is becoming automated, data-driven, and AI-enabled.
- AI supports prediction, planning, control, and traffic management
- Safety-critical autonomy also needs:
  - quantified uncertainty and verifiable safety guarantees



Thesis: Safe aerospace autonomy requires uncertainty-aware AI, not just AI-enabled optimisation

Why Is This Hard? Data, Validation, and Rare Events



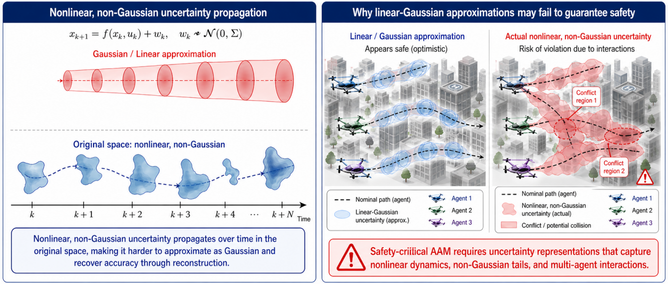
Safe autonomy requires uncertainty-aware evidence and decision making.

- Aerospace needs both evidence generation and uncertainty-aware validation

The Technical Core

$$x_{k+1} = f(x_k, u_k) + w_k, \qquad w_k \not\sim \mathcal{N}(0, \Sigma)$$

- **Goal:** enforce safety constraints despite uncertain dynamics
- **Difficulty:** uncertainty is shaped by nonlinear dynamics and rarely remains Gaussian
- **Risk:** linear–Gaussian approximations can underestimate safety violations
- **Need:** uncertainty-aware safety filters with verifiable guarantees



Core question: How can safety control remain reliable when uncertainty is nonlinear and non-Gaussian?

Talk Roadmap: CBFs for Uncertainty-Aware Autonomy

Motivation
Why aerospace autonomy needs CBFs beyond nominal safety constraints

$\Downarrow$

Part I: Spectral Probabilistic CBF
Foundation: probabilistic safety under nonlinear, non-Gaussian uncertainty

$\Downarrow$

Part II: Adaptive Consensus Safety Filter
Multi-Agent Case: multi-UAV safety under delayed communication

$\Downarrow$

Takeaway
Toward real-time, uncertainty-aware CBFs for assured aerospace autonomy

Outline

Theoretical Challenge: Probabilistic Safe Control

$$\begin{aligned} \min_{u_k \in \mathcal{U}} \quad & \|u_k - u_{\text{nom}}\|^2 \\ \text{s.t.} \quad & x_{k+1} = f(x_k, u_k) + w_k, \quad w_k \sim \mathcal{P} \\ & \mathbb{P}\big(h(x_{k+1}) \geq 0 \mid x_k, u_k\big) \geq 1 - \delta \quad \text{(Probabilistic Forward Invariance)} \end{aligned}$$

- Gaussian assumption are invalid
- This is a probabilistic CBF condition, not merely a static chance constraint
- Nonlinear dynamics and non-Gaussian noise distort future state distributions
- Evaluating the probability requires integration over nonlinear safe-evolution sets

Core difficulty: Probabilistic forward invariance under nonlinear dynamics and arbitrary noise is generally non-convex and computationally intractable

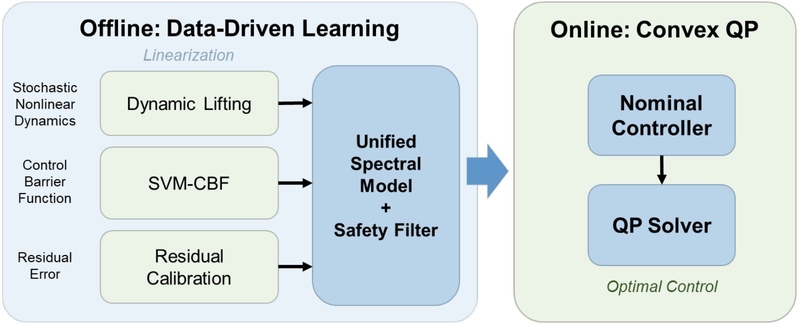
Related Works

Approach	Key Advantage	Primary Limitation
Probabilistic CBF	Rigorous stochastic safety bounds	Relies on known dynamics & Gaussian noise assumptions
Data-Driven CBF	Model-free synthesis of complex safe sets	Lacks robust guarantees under stochastic transitions

- Dynamic systems require safety despite nonlinearities and complex disturbances.
- The control algorithm must remain computationally efficient for fast online decision-making.

Need: A distribution-aware and computationally tractable probabilistic CBF-QP formulation.

Proposed Framework: Spectral CBF



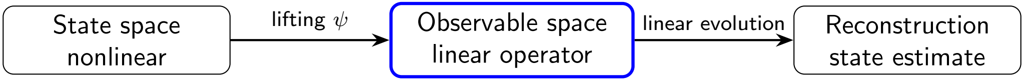
- **Offline:** map nonlinear, non-Gaussian safety boundaries into a convex spectral distance
- **Online:** solve a convex QP with an empirical margin

Main idea: transform dynamics and safety constraints into a tractable spectral representation, then solve probabilistic CBF-QP online

Offline: Dynamics Lifting via Koopman/RFF Features

$$x_{k+1} = F(x_k) \qquad \text{nonlinear dynamics in state space}$$

$$\mathcal{K}\psi(x) = \psi(F(x))$$



- Koopman acts on observables, not directly on states
- Nonlinear dynamics become linear in an infinite-dimensional function space
- Practical methods approximate this space with finite-dimensional features

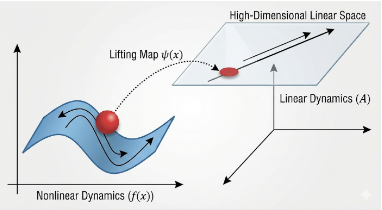
Offline: Dynamics Lifting via Koopman/RFF Features

$$x_{k+1} = f(x_k, u_k) + w_k \qquad \Longrightarrow \qquad \mu_{k+1} = A_\psi \mu_k + B_\psi u_k$$

$$\psi(x) = \begin{bmatrix} 1 & x^\top & \zeta_{\text{dyn}}(x)^\top & \zeta_{\text{con}}(x)^\top \end{bmatrix}^\top, \qquad \mu_k = \mathbb{E}[\psi(x_k)]$$

State x_k —lifting→ RFF / Koopman features $\psi(x_k)$ —expectation→ Distribution embedding $\mu_k = \mathbb{E}[\psi(x_k)]$ —prediction→ Linear lifted prediction $\mu_{k+1} = A_\psi \mu_k + B_\psi u_k$

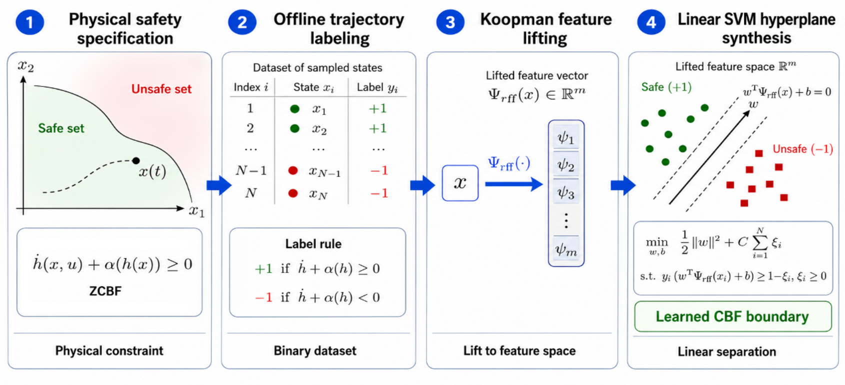
- Random Fourier Features generate finite-dimensional spectral observables
- EDMDc[a] learns A_ψ, B_ψ from data
- Nonlinear stochastic prediction becomes linear prediction in lifted feature space



[a] EDMD: Extended Dynamic Mode Decomposition

Offline: Spectral CBF Synthesis via SVM

- Embeds physical safe set into the same feature space (Ψ) used by the Koopman dynamics.



Learning the geometry of the CBF-valid set in the chosen feature space

$\rightarrow$ Complex and non-convex safe sets are transformed into deterministic linear constraints

Offline: Calibration Margin

- **SVM Residual (m_{svm}):**
  The maximum distance of false(unsafe) data.

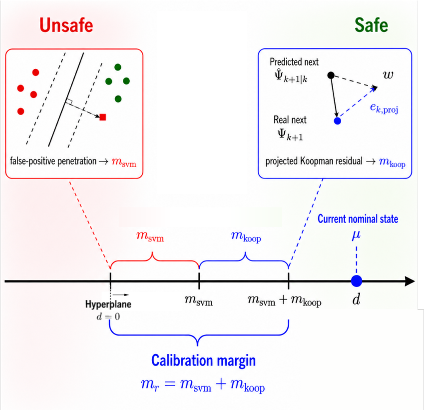
$$m_{\text{svm}} = \max_{x \in \mathcal{D}_{\text{unsafe}}} \left\{ 0,\ w^\top \Psi_{\text{rff}}(x) + b \right\}$$

- **Dynamic Residual(m_{koop}):**
  The statistical trajectory-wise prediction error over the calibration dataset.

$$m_{\text{koop}} = \text{Quantile}_{1-\delta}\left(\left\{ \max_{k \in \tau_i} \left(0,\ w^\top \left(\hat{\Psi}_{err}^{(i,k)} \right) \right) \right\}_{i \in \mathcal{D}_{\text{cal}}} \right)$$

- **Unified Calibartion Margin:**

$$m_r = m_{\text{svm}} + m_{\text{koop}}$$



Online: Convex Spectral CBF-QP

1. **Enforcing Probabilistic Safety in Koopman Space**
   Applying the expectation operator over the future state distribution

$$\mathbb{E}\left[w^\top \Psi_{\text{rff}}(x_{k+1}) + b\right] \geq m_r \quad \Longrightarrow \quad w^\top \mu_{k+1} + b \geq m_r$$

2. **Resulting Spectral CBF-QP Formulation**
   Using Koopman dynamics ($\mu_{k+1} = A\mu_k + Bu_k$), the condition becomes an affine constraint with respect to u_k

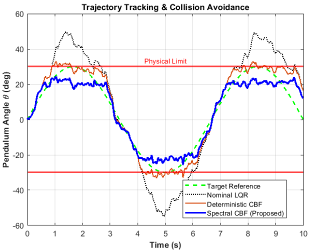
$$\begin{aligned} \min_{u_k} \quad & \|u_k - u_{\text{nom}}\|^2 \\ \text{s.t.} \quad & w^\top A\mu_k + w^\top Bu_k + b - m_r \geq 0 \end{aligned}$$

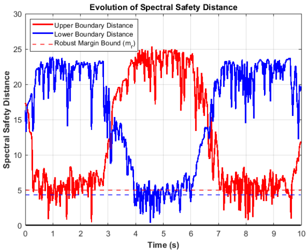
Message: Probabilistic safe control problem turns into a convex QP structure.

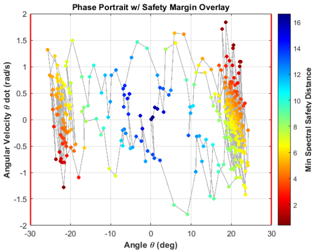
KAIST AIS

Simulation, Results, and Analysis

Setup: Inverted Pendulum under non-Gaussian noise (Nominal: Discrete LQR)







Takeaway: Spectral-CBF enforces safety constraints under non-Gaussian dynamics with computational efficiency (Avg. online time: 0.36 ms).

Contributions & Discussions

Theorem 1: Probabilistic Forward-Invariance Guarantee

Given a risk budget $\delta_k \in (0,1)$, if the control input satisfies the calibrated spectral safety constraint,

$$w^\top (A\Psi_{\text{rff}}(x_k) + Bu_k) + b \geq m_r,$$

then the probabilistic CBF condition is satisfied:

$$\mathbb{P}(h(x_{k+1}) - (1-\gamma)h(x_k) \geq 0 \,|\, x_k, u_k) \geq 1 - \delta_k$$

Consequently, if $h(x_k) \geq 0$, then

$$\mathbb{P}(h(x_{k+1}) \geq 0 \,|\, x_k, u_k) \geq 1 - \delta_k$$

- **Key contributions**
  - Learns a spectral surrogate of the CBF-valid safe-evolution condition
  - Converts non-Gaussian probabilistic invariance into a deterministic spectral constraint
  - Combines SVM boundary error and Koopman prediction error into one calibrated margin m_r
  - Enables real-time safety filtering through an affine QP constraint

Outline

Fundamental Challenge: Chance-Constrained Safety

$$\begin{aligned} \min_{\mathbf{u}_i} \quad & \frac{1}{2}\left\|\mathbf{u}_i - \mathbf{u}_{i,\text{nom}}\right\|^2 \\ \text{s.t.} \quad & \mathbf{u}_i \in \mathcal{U} \\ & \mathbb{P}\big(\|\mathbf{p}_i(t) - \mathbf{p}_j(t)\| \geq \delta_{\text{safe}}\big) \geq 1-\alpha \quad \forall j \in \mathcal{N}_i \end{aligned}$$

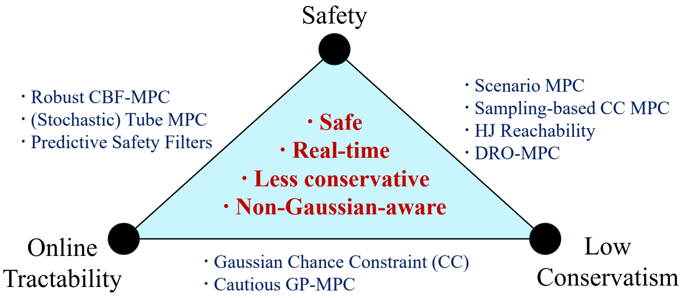
$\Rightarrow$ Pairwise Chance Constraints

- Non-convex pairwise chance constraints → Analytically intractable
- Time-varying & stochastic communication delays → Non-Gaussian uncertainty
- Information asymmetry from stale states → Inconsistent local safety sets
- Naive worst-case robustification → Conservatism

Core Question

How can we guarantee multi-UAV safety under non-Gaussian uncertainty induced by communication delays, while reducing conservatism?

Related Works: The Trilemma of Multi-Agent Safety



- Communication delays induce heavy-tailed, non-Gaussian uncertainty
- The controller must remain fast enough for online decision making

Need: Less conservative RCBFs (Safety Filter) for Multi-Agent Systems with real-time safety guarantees

Problem Formulation: Insight

Goal: guarantee true pairwise safety

$$\mathbb{P}(\|p_i(t) - p_j(t)\| \geq \delta_{\text{safe}}) \geq 1 - \alpha$$

Challenge: $p_j(t)$ is unknown

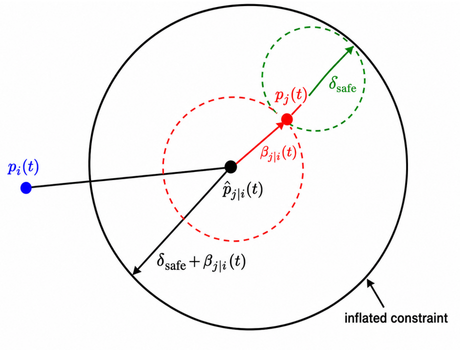
UQ condition:

$$\mathbb{P}\left(\|p_j(t) - \hat{p}_{j|i}(t)\| \leq \beta_{j|i}(t)\right) \geq 1 - \alpha$$

Controller condition:

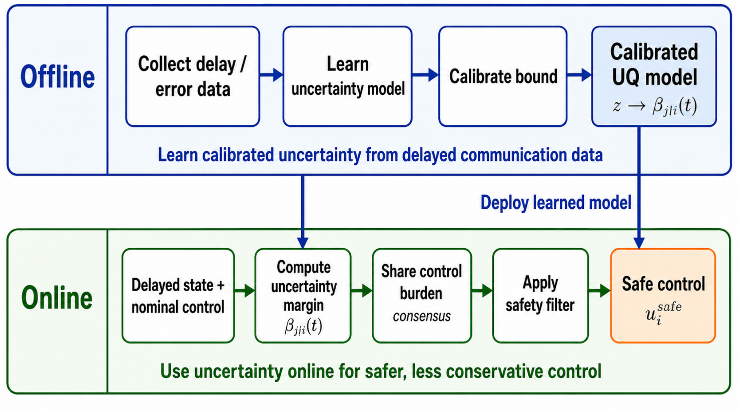
$$\|p_i(t) - \hat{p}_{j|i}(t)\| \geq \delta_{\text{safe}} + \beta_{j|i}(t)$$

$$\Longrightarrow \mathbb{P}(\|p_i(t) - p_j(t)\| \geq \delta_{\text{safe}}) \geq 1 - \alpha$$



How to reduce this conservatism without sacrificing safety?

Proposed Framework: Adaptive Consensus Safety Filter

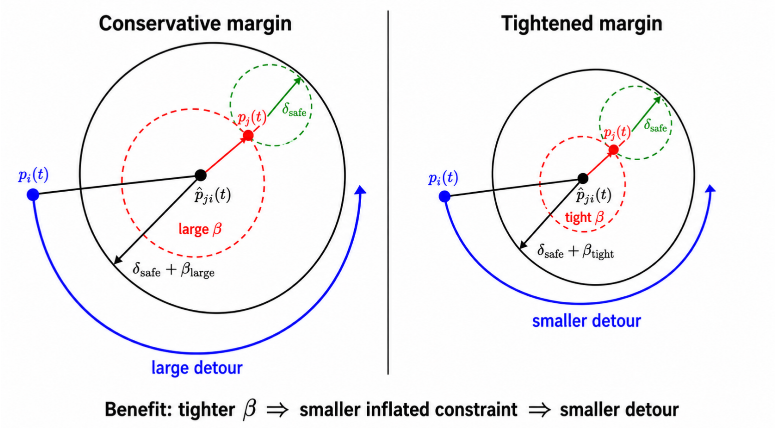


Main idea: Translate communication delay into spatial uncertainty bounds, then reduce conservatism via consensus-based burden sharing.

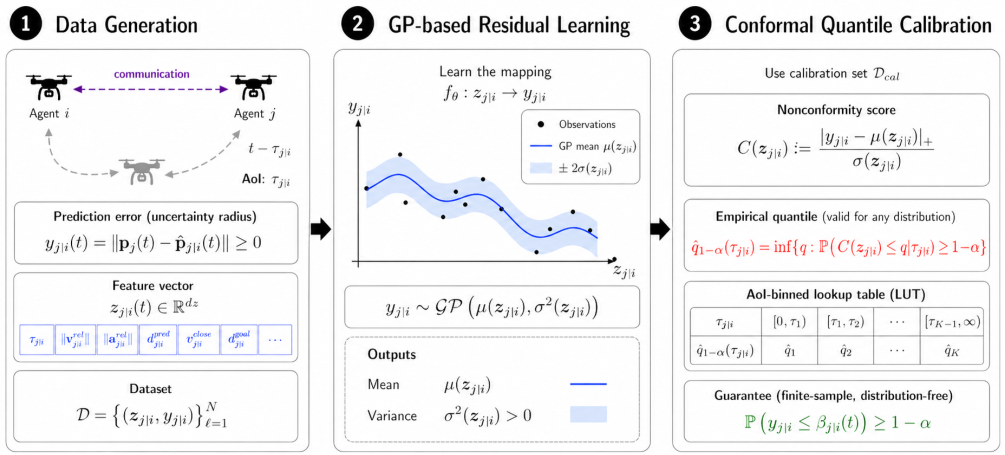
Stage 1: Data-Driven Learning and Statistical Calibration

Aim: learn a valid but tighter uncertainty margin $\beta_{j|i}(t)$ that satisfies

$$\mathbb{P}\big(\|p_j(t) - \hat{p}_{j|i}(t)\| \le \beta_{j|i}(t)\big) \ge 1 - \alpha.$$

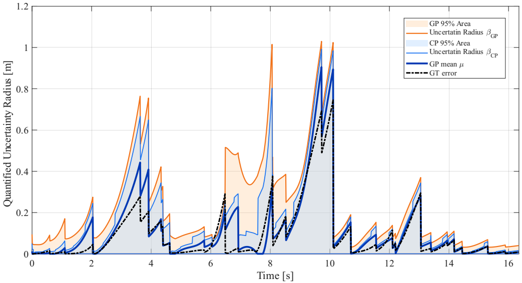


Stage 1: Data-Driven Learning and Statistical Calibration



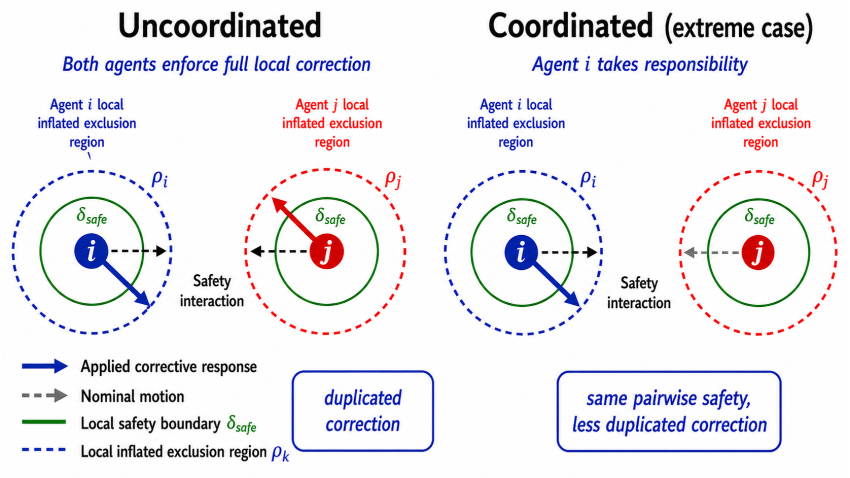
Stage 1: Empirical Margin Calibration via CP

- **GP-only UQ**
  - Tail uncertainty is handled through a Gaussian quantile, which can inflate the margin.
  - $\beta_{\mathrm{GP}} = \mu + q^{\mathrm{GP}}_{1-\alpha}\sigma$ (Gaussian-tail)
- **CP-calibrated UQ**
  - Use an empirical conformal quantile, yielding tighter bounds.
  - $\beta_{\mathrm{CP}} = \mu + \hat{q}^{\mathrm{CP}}_{1-\alpha}\sigma$ (data-calibrated)



Message: Stage 1 replaces the Gaussian tail quantile with an empirical conformal quantile to obtain a tighter deterministic margin β

Stage 2: Allocation of Safety Responsibility



Stage 2: Time-Varying Safe Set via HO-RCBF

$$\underbrace{\|\mathbf{p}_i(t) - \hat{\mathbf{p}}_{j|i}(t)\| \geq \delta_{\text{safe}} + \beta_{j|i}(t)}_{\text{Deterministic Safety Constraint}} \xrightarrow{\text{calibrated UQ}} \underbrace{\mathbb{P}\Big(\|\mathbf{p}_i(t) - \mathbf{p}_j(t)\| \geq \delta_{\text{safe}}\Big) \geq 1 - \alpha}_{\text{Original CC (Goal)}}$$

Definition 1: Time-Varying Safe Set

To guarantee safety, the system state must remain within the safe set $\mathcal{C}_{ij}$, defined by a high-order robust control barrier function (HO-RCBF) h_{ij}:

$$h_{ij}(\mathbf{x}_i, \hat{\mathbf{x}}_{j|i}, t) = \|\mathbf{p}_i(t) - \hat{\mathbf{p}}_{j|i}(t)\|^2 - (\delta_{\text{safe}} + \beta_{j|i}(t))^2$$

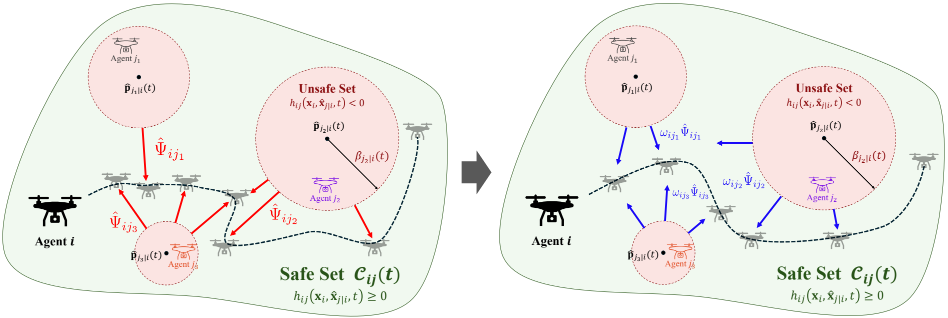
$$\mathcal{C}_{ij}(t) = \Big\{\mathbf{x}_i \in \mathbb{R}^n \;\Big|\; h_{ij}(\mathbf{x}_i, \hat{\mathbf{x}}_{j|i}, t) \geq 0\Big\}$$

$$\psi_4 \geq 0 \text{ (HO-RCBF cdn.)} \xrightarrow{\textbf{control-affine safety constraints}} \hat{\mathbf{G}}_{ij}^\top \mathbf{u}_i \leq \hat{\Psi}_{ij}$$

$$\Big(\psi_0 = h_{ij}, \qquad \psi_\ell = \dot{\psi}_{\ell-1} + \alpha_\ell(\psi_{\ell-1}), \quad \alpha_\ell(s) = \gamma_\ell s, \quad \ell = 1, \ldots, 4\Big)$$

Stage 2: Uncoordinated Safety Constraints

$$\hat{\mathbf{G}}_{ij}^{\top}\mathbf{u}_i \le \hat{\Psi}_{ij}(\beta_{j|i}) = \overbrace{\Psi_{ij} - \underbrace{\mathcal{M}_i(\beta_{j|i},\gamma_i)}_{\text{Conservative Margin}} + \underbrace{\mathcal{K}_i(\gamma_i)}_{\text{Kinematic Error}}}^{\text{Control-Effective Safety Budget}}$$



Each agent enforces the full $\hat{\Psi}_{ij}$ independently, leading to conservatism

Stage 2: Coordinated Safety Constraints

Uncoordinated enforcement

$$\hat{G}_{ij}^{\top} u_i \leq \hat{\Psi}_{ij}$$

$$\hat{G}_{ji}^{\top} u_j \leq \hat{\Psi}_{ji}$$

- Each agent reacts to its full local safety demand
- Conservative due to duplicated correction

Coordinated enforcement

$$\hat{G}_{ij}^{\top} u_i \leq \omega_{ij} \hat{\Psi}_{ij}$$

$$\hat{G}_{ji}^{\top} u_j \leq \omega_{ji} \hat{\Psi}_{ji}$$

- ω_{ij} allocates agent i's responsibility
- Corrective burden is shared across the pair

$$\omega_{ij} + \omega_{ji} = 1$$

$\hat{\Psi}_{ij}$: required local corrective safety budget

Key idea: Same safety goal, allocated correction

Stage 2: Control Burden Allocation via C-ADMM

- Centralized Formulation

$$\begin{aligned}\min_{\mathbf{u}_i,\omega_{ij}} \quad & \sum_{i=1}^{N}\left(\frac{1}{2}\|\mathbf{u}_i-\mathbf{u}_{i,\text{nom}}\|^2+\sum_{j\in\mathcal{N}_i}\frac{\eta_\omega}{2}(\omega_{ij}-\omega_{ij,d})^2\right)\\ \text{s.t.} \quad & \hat{\mathbf{G}}_{ij}^{\top}\mathbf{u}_i \le \omega_{ij}\hat{\Psi}_{ij},\ \forall j\in\mathcal{N}_i\\ & \omega_{ij}+\omega_{ji}=1,\ \forall j\in\mathcal{N}_i \quad \Rightarrow \text{Coupled-Constraint}\\ & \mathbf{u}_i\in\mathcal{U}\end{aligned}$$

Scaled Augmented Lagrangian

$$\mathcal{L}_{i,\rho}(\mathbf{u}_i,\omega_{ij},\tilde{\omega}_{ij},\bar{\lambda}_{ij}) = \frac{1}{2}\|\mathbf{u}_i-\mathbf{u}_{i,\text{nom}}\|^2+\sum_{j\in\mathcal{N}_i}\left(\frac{\eta_\omega}{2}(\omega_{ij}-\omega_{ij,d})^2+\frac{\rho}{2}(\omega_{ij}-\tilde{\omega}_{ij}+\bar{\lambda}_{ij})^2-\frac{\rho}{2}\bar{\lambda}_{ij}^2\right)$$

- $\tilde{\omega}_{ij}\in\Omega=\{\tilde{\omega}\mid\tilde{\omega}_{ij}+\tilde{\omega}_{ji}=1\}$: Consensus variable.
- $\bar{\lambda}_{ij}=\lambda_{ij}/\rho$: Scaled dual variable.
- η_ω,ρ: Weighting & penalty params.

Stage 2: Control Burden Allocation via C-ADMM

- Distributed 3-step updates
  1. **Primal Update (Local QP)**

$$(\mathbf{u}_i^{k+1}, \omega_{ij}^{k+1}) = \arg\min_{\mathbf{u}_i, \omega_{ij}} \mathcal{L}_{i,\rho}(\mathbf{u}_i, \omega_{ij}, \tilde{\omega}_{ij}^k, \bar{\lambda}_{ij}^k)$$

$$\text{s.t. } \hat{\mathbf{G}}_{ij}^\top \mathbf{u}_i \leq \omega_{ij} \hat{\Psi}_{ij}$$

$$\mathbf{u}_i \in \mathcal{U}$$

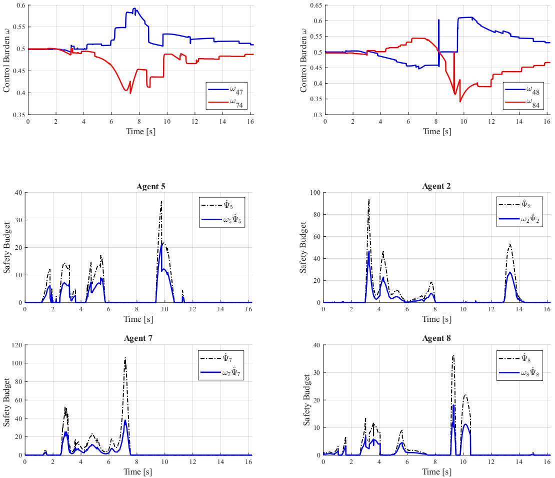
  2. **Consensus Update**

$$\tilde{\omega}_i^{k+1} = \arg\min_{\tilde{\omega}_i} \mathcal{L}_{i,\rho}(\mathbf{u}_i^{k+1}, \omega_{ij}^{k+1}, \tilde{\omega}_{ij}, \bar{\lambda}_{ij}^k)$$

$$\text{s.t. } \tilde{\omega}_{ij} + \tilde{\omega}_{ji} = 1, \ \tilde{\omega}_{ij} \in [\underline{\omega}, \bar{\omega}]$$

  3. **Dual Update**

$$\bar{\lambda}_{ij}^{k+1} = \bar{\lambda}_{ij}^k + (\omega_{ij}^{k+1} - \tilde{\omega}_{ij}^{k+1})$$

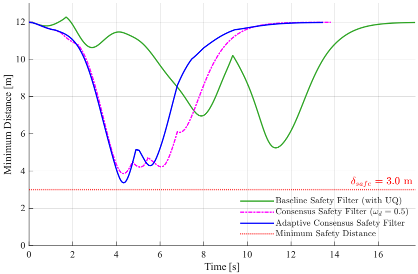


Simulation Results and Analysis

Setup: 8-UAV antipodal swapping, **Comm.**: 5 Hz, **Max delay**: 2.0 s

- **Takeaway**: Conservatism is reduced not by sacrificing safety, but by redistributing it.
  - Safety preserved
  - Online tractability
  - Robust to non-Gaussian uncertainty
  - Reduced conservatism

under unreliable communication!



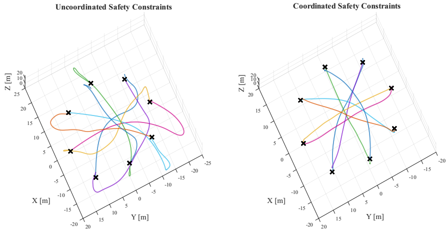
Method	Avg. Flight Distance [m]	Arrival Time [s]	Avg. Compute Time per Step [ms]	Control XY-Effort ($\sum \|\mathbf{u}_{xy}\|$) [rad]	Control Z-Effort ($\sum \|\mathbf{u}_z\|$) [m/s^2]	Control Smoothness ($\sum \|\Delta \mathbf{u}\|$) [n.u.]
Baseline Safety Filter (with UQ)	54.54	17.69	**0.5369**	659.41	24791.27	52.097
Consensus Safety Filter ($\omega = 0.5$)	41.63	13.83	0.6694	414.54	19381.75	49.562
Adaptive Consensus Safety Filter	**40.55**	**13.47**	0.6619	**392.88**	**18877.24**	**41.440**

Is Safety Still Guaranteed?

Theorem 2: Safety Certificate under Adaptive Consensus

$$\underbrace{\mathcal{M}_i(\beta_{j|i}, \gamma_i) - \mathcal{K}_i(\gamma_i)}_{\text{net robustness reserve}} \geq \underbrace{\frac{\bar{u}\|\Delta_{\mathbf{G}_{ij}}\|}{\omega_{ij}}}_{\substack{\text{model-mismatch penalty} \\ \text{under allocation}}} \qquad \forall i, j$$

- $\beta_{j|i}$ creates robustness reserve against delayed-state uncertainty
- $\mathcal{K}_i$ captures relative-motion CBF residuals
- Smaller ω_{ij} reduces local burden, but tightens certification
- Sufficient condition: no certificate does not imply unsafe



Implication: Allocate correction within available robustness reserve

Outline

Conclusions

- Presented two uncertainty-aware CBF developments toward aerospace autonomy
- **Probabilistic CBF foundation:** spectral formulation for nonlinear, non-Gaussian uncertainty
- **Multi-agent CBF extension:** adaptive consensus safety filters under delayed communication
- Both translate uncertainty-aware safety requirements into tractable online control constraints
- Results indicate reduced conservatism while retaining formal safety certificates

Takeaway: Uncertainty-aware CBFs provide a pathway from probabilistic safety foundations to cooperative autonomy

Future Works

- Improve scalability of spectral uncertainty representations
- Apply the spectral CBF to practical Aerospace applications
- Extend probabilistic safety guarantees under rare events and distribution shift
- Develop adaptive tuning laws for γ_i, ρ, and allocation weights
- Analyze large-scale behavior for dense UAV / AAM traffic networks
- Validate using hardware experiments and high-fidelity aerospace simulation

Long-term goal: Real-time, certifiable CBF safety filters for safety-critical aerospace autonomy

Thank You!