www.ijarp.com ISSN 2456-9992 Page: 01-33

International Journal Advanced Research Publications

Review Article

Volume: 02
Issue: 07

FEASIBILITY-AWARE GEOSPATIAL PATH PLANNING AND RIS-ASSISTED COMMUNICATION FOR ENERGY-CONSTRAINED MULTI-UAV SYSTEMS

Muhammad Rashid Majeed, Mst Jannatul Kobra*

Computer Science , Nanjing University of Information Science and Technology, 219, Ningliu Road, Pukou, Nanjing, Jiangsu.

Article Received: 25 May 2026, Article Revised: 13 June 2026, Published on: 03 July 2026

***Corresponding Author: Mst Jannatul Kobra**

Computer Science , Nanjing University of Information Science and Technology, 219, Ningliu Road, Pukou, Nanjing, Jiangsu.

Doi: https://doi-doi.org/101555/ijarp.6679

ABSTRACT

Autonomous multi-UAV systems are increasingly deployed in applications such as disaster response, environmental monitoring, and smart city management, where dynamic and uncertain environments pose significant operational challenges. In such situations, UAVs have to cope with geospatial risks, establish reliable communication links, and strictly enforce constraints on energy and latency requirements. Conventional strategies usually tackle these challenges individually, making it hard to ensure feasible mission performance in uncertain environmental settings. To address this research gap, this paper proposes a feasibility-aware co-design framework that integrates geospatial risk-aware path planning, reconfigurable intelligent surface (RIS)-assisted wireless communication, and adaptive multi-agent decision-making within a unified control system. In contrast, the novel strategy utilizes risk-oriented geospatial planning, designs efficient phase configurations for RISs to improve communication conditions in the presence of uncertainties, and enforces constraints on energy and latency through a dual-primal control architecture. Moreover, a multi-agent reinforcement learning scheme helps make adaptive decisions for such a system without the need for manual reparameterization. Simulations in stochastic channel conditions and dynamic obstacle environments demonstrate that the framework improves spectral efficiency by approximately 30% compared to non-RIS baselines and reduces energy consumption by over 20% relative to non-geospatial strategies.The framework achieves high feasibility across all simulated scenarios. These results demonstrate the applicability of integrating geospatial planning, communi-

cation management, and feasibility control in multi-UAV missions. The proposed framework provides a scalable foundation for future autonomous UAV systems, enabling energy-efficient, reliable, and safe operation in complex real-world environments.

KEYWORDS: Autonomous multi-UAV systems, feasibility-aware control, reconfigurable intelligent surfaces (RIS), adaptive multi-agent learning, robust wireless communication.

1. INTRODUCTION

The deployment of autonomous unmanned aerial vehicles (UAVs) in smart cities, disaster recovery missions, environmental monitoring, and future wireless communication networks poses many challenges due to their dynamism and unpredictability [1]–[4]. In contrast to traditional single-UAV systems, multi-UAV networks must simultaneously tackle challenges in navigation safety, communication reliability, and harsh resource constraints, where energy exhaustion and latency constraint violations can severely impact mission success [5]–[7]. Conventional approaches to resolving these problems do not consider them concurrently and cannot provide any feasibility assurance under environmental uncertainty [8], [9]. Traditional path planning algorithms and control heuristics, although computationally efficient, have generally overlooked communication-aware decision-making and lacked explicit feasibility analysis for coupled energy-latency constraints [10]–[14]. Reinforcement learning (RL) and multi-agent reinforcement learning (MARL) algorithms have been proposed to enhance the adaptability of UAVs. Still, training instability, non-intuitive solutions, and tight constraints on satisfaction make them impractical in real-world scenarios [15]–[18]. Concurrently, the concept of reconfigurable intelligent surfaces (RIS) has also been identified as a promising solution for enhancing the spectral efficiency and robustness of UAV-assisted wireless communication systems, but current studies on RIS optimization are generally conducted in a manner that is decoupled from navigation and feasibility analysis [19]–[21]. As such, a gap in research exists regarding a holistic approach to combining the geospatial understanding, communication optimization via the use of RIS technology, and enforceability of feasibility [22], [23]. To fill this research void, this paper presents a feasibility-aware co-design framework that closely integrates geospatial risk-aware path planning with RIS-assisted wireless communication and the explicit enforcement of energy-latency constraints via a primal-dual control approach, while also incorporating an adaptive multi-agent decision layer for online modification in response to environmental perturbations [24]. This paper summarizes the overall outline of the paper and describes the contents of each section, including the methodological approach, simulations, and assessment metrics like spectral efficiency, energy consumption, latency, and feasibility [25]–[27]. The simulation results show that the proposed framework can improve the communication performance and energy efficiency while ensuring the feasibility satisfaction, which provides a reliable basis for the future autonomous multi-UAV communication system [28]–[30].

1.1 Novelty

The core novelty of the present study is the combined use of geospatial risk-aware path planning, radio-interference suppression (RIS)-based wireless communications, and multi-agent reinforcement learning (MARL). As opposed to current literature, which explores each of these factors individually, the present study considers all of these factors concurrently using a primal-dual control methodology and online adaptation through MARL. The proposed solution thus guarantees that the trajectory followed by the UAV will be feasible while ensuring efficient communication.

2. Related Work

Existing works on autonomous UAV systems cover a range of related areas, such as geospatial path planning, wireless communication optimization, and learning-based control. Traditional UAV navigation systems mainly focus on deterministic or heuristic planners, like graph-based and sampling-based planners, which are computationally efficient but lack communication awareness and feasibility [31]–[34]. To deal with the uncertainty and adaptability, reinforcement learning and deep reinforcement learning have been increasingly applied to UAV navigation and coordination; however, these approaches tend to optimize performance metrics without considering the feasibility of energy or latency constraints in safety-critical applications [35]–[39]. Multi-agent reinforcement learning (MARL) has been further developed to support cooperative UAV networks, which can facilitate decentralized decision-making and scalability, but existing MARL-based methods are prone to unstable convergence, hyperparameter sensitivity, and poor robustness against environmental disturbances [40]–[44]. Meanwhile, energy-efficient UAV control and trajectory optimization have been explored to maximize mission lifetimes, but most existing works on energy-efficient UAV control and trajectory optimization rely on simplified communication models or static environments, which neglect the intertwined relationship between mobility and wireless link quality [45]–[48]. Reconfigurable intelligent surfaces (RIS) have recently been introduced as a promising tool for improving the spectral efficiency and coverage of UAV-assisted wireless communications, with many studies showing significant communication performance improvements through adaptive phase control; however, these studies have mostly separated RIS optimization from UAV trajectory planning and failed to consider system-level feasibility issues [49]–[53]. Some efforts have been made to investigate joint communication and trajectory planning for UAVs, but these studies usually involve centralized optimization with perfect channel knowledge or lack the ability to cope with real-time system disturbances [54]–[58]. More recently, some studies have started to investigate integrated solutions that combine learning, communication, and control; however, these studies usually aim at maximizing performance rather than satisfying strict feasibility constraints, and the scalability problem of multi-UAV systems with guaranteed feasibility remains largely open [59]–[62]. Therefore, despite the significant progress made in individual building blocks, an integrated feasibility-aware co-design solution that combines geospatial risk, RIS-assisted communication, and adaptive multi-agent decision-making with practical feasibility constraints is still absent in the current literature [63]–[65].

2.1 Problem Formulation

We formulate a cooperative multi-UAV system in a dynamic environment with time-varying geospatial obstacles and stochastic wireless channels. Each UAV needs to navigate from a pre-specified start point to a target destination while ensuring continuous communication with the ground infrastructure via a reconfigurable intelligent surface (RIS). The system state encodes both navigation risk and communication quality, as well as information related to constraints, while control variables represent motion risk, risk tolerance in route selection, and communication power control. The aim is to maximize the long-term system utility measuring spectral efficiency and robustness of operation, under very tight energy and latency constraints for each mission. Energy use is a function of route distance, traversal speed, and geospatial risk, while communication latency is affected by channel conditions and RIS phase settings. To guarantee safe and reliable system operation, feasibility constraints are enforced explicitly by incorporating the cost of violating energy and latency constraints, thus ensuring satisfaction of the constraints over the entire mission horizon. The problem is cast as a constrained sequential decision-making problem, where system dynamics are driven by environmental disturbances and decentralized decision-making policies, and the task is to optimize simultaneously navigation and communication system performance simultaneously while satisfying feasibility constraints.

3. METHODOLOGY

This section describes the methodology for designing an efficient system for autonomous UAV navigation, incorporating geospatial risk-aware planning, RIS-assisted communication, and adaptive control using Adaptive Multi-Agent Reinforcement Learning (MARL). The aim is to guarantee real-time feasibility and energy efficiency while optimizing communication performance.

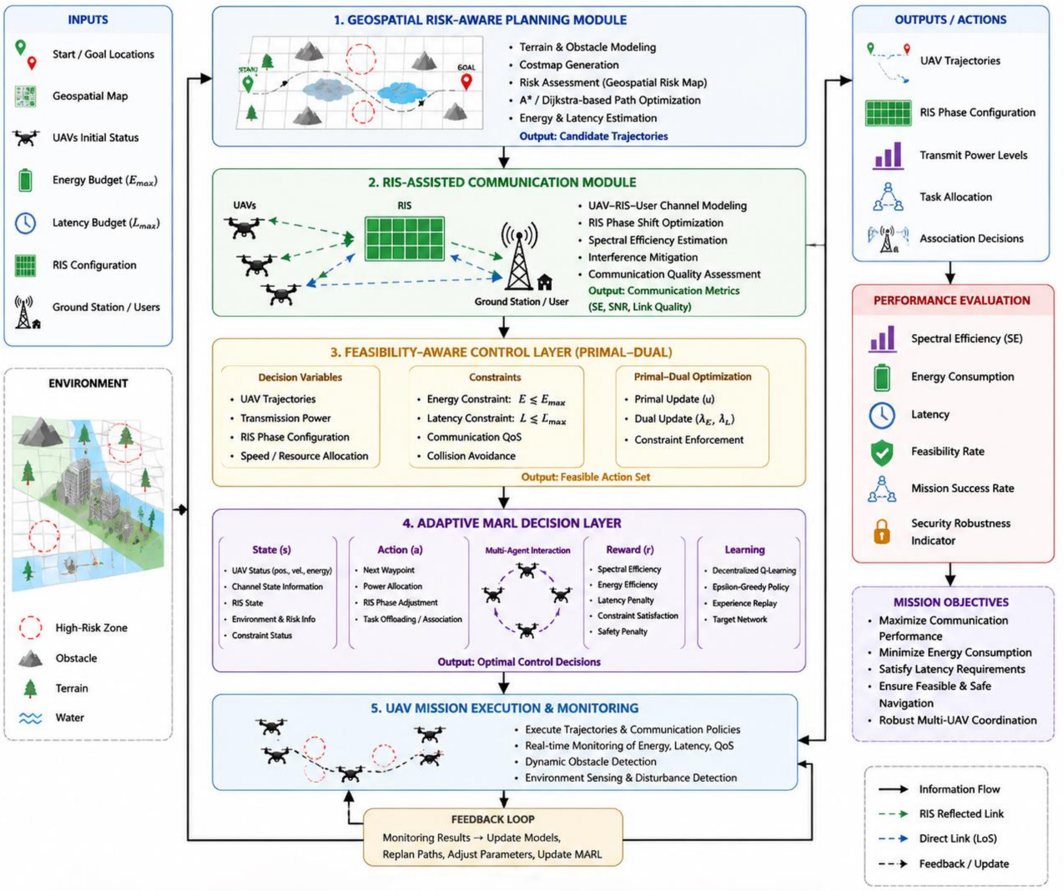

Fig1. Proposed architecture of the feasibility-aware multi-UAV framework.

The architecture (Fig 1) integrates geospatial risk-aware path planning, RIS-assisted communication optimization, feasibility-aware primal-dual control, and adaptive multi-agent reinforcement learning (MARL) into a unified decision-making framework. Candidate trajectories generated by the geospatial planner are evaluated through RIS-assisted communication analysis and feasibility-aware control before being processed by the MARL decision layer. The selected actions are executed by the UAV network, while continuous monitoring and feedback enable real-time adaptation to dynamic environments under energy and latency constraints.

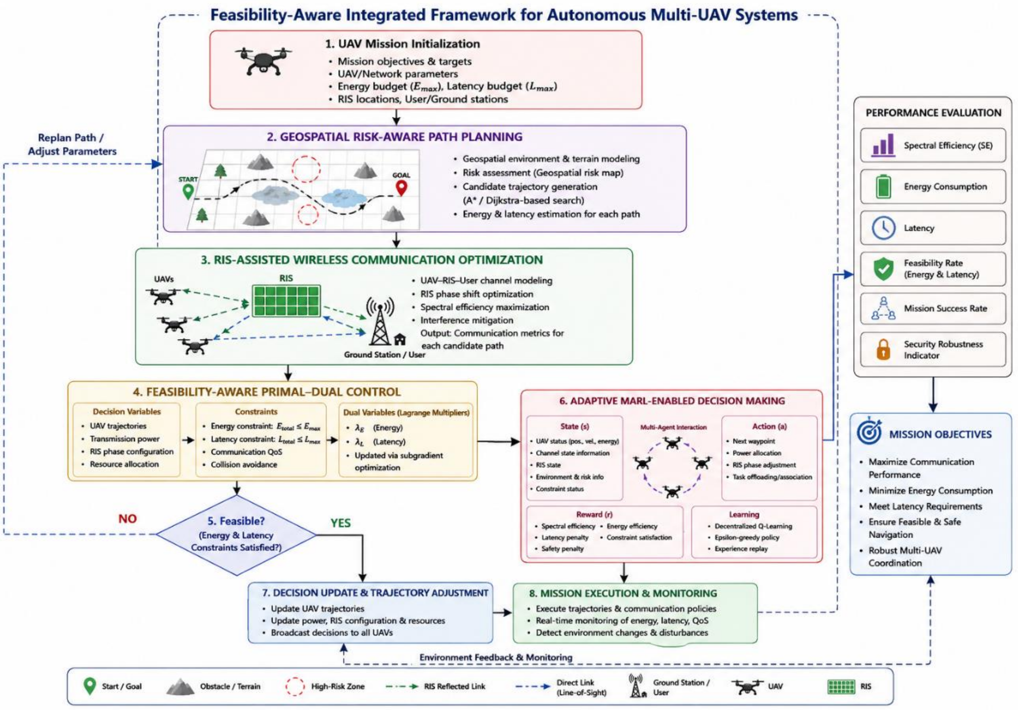

Fig 2. Flowchart of the proposed multi-UAV framework.

The following diagram provides a visual representation of this process, highlighting the relationship among geospatial risk-aware UAV path planning, RIS-enabled communications, feasibility-aware UAV control, and adaptive MARL-enabled decision making during UAV operations.

3.1Geospatial Risk-Aware Planning

In UAV navigation, geospatial risk-aware planning is the process of finding safe and optimal paths by considering the terrain roughness and the presence of obstacles. This section explains how the environment is represented and how the optimal path is calculated.

Algorithm 1 outlines the integrated geospatial risk-aware path planning and feasibility control workflow

Input:

G ← Geospatial grid map

C ← Costmap (terrain + obstacle weighted)

s_start ← Start position

s_goal ← Goal position

E_max ← Energy budget

L_max ← Latency budget

Output:

P_opt ← Optimal feasible path

Feas_flag ← Feasibility status

Procedure

1. Initialize empty path: P_opt ←∅
2. Generate costmap C from geospatial grid G
3. Define feasible region M from obstacle-free cells
4. Apply graph-based search (A* / Dijkstra) on C within M
5. Generate candidate path P_cand
6. Estimate energy consumption E for P_cand
7. Estimate latency L for P_cand
8. If E ≤ E_max AND L ≤ L_max then
9. Set Feas_flag ← TRUE
10. Set P_opt ← P_cand
11. Else
12. Adjust path cost weights (increase obstacle/terrain penalty)
13. Recompute P_cand
14. Re-evaluate E and L
15. If constraints are satisfied, then
16. Set Feas_flag ← TRUE
17. Set P_opt ← P_cand
18. Else
19. Set Feas_flag ← FALSE
20. End If
21. Return (P_opt, Feas_flag)

The feasibility control layer performs only a single replanning attempt when the initially generated trajectory fails to satisfy the predefined energy and latency constraints. During this process, the path-planning cost weights are adjusted to generate an alternative trajectory. The revised trajectory is subsequently evaluated against all feasibility requirements. If the updated trajectory satisfies the specified constraints, it is accepted for execution. Otherwise, the mission is classified as infeasible, and the planning process is terminated without further replanning attempts. This bounded replanning strategy ensures predictable computational complexity while maintaining real-time operational efficiency.

3.1.1 Costmap Generation

The environment is represented as a grid, in which every cell (x,y) is associated with a cost of traversal. The cost map C(x,y) is updated dynamically according to the UAV's location.

$$C(x,y) = \text{Terrain}(x,y) + \alpha \cdot \text{Obstacle}(x,y) \tag{1}$$

3.1.2 Path Optimization

The path of the UAV is determined to be the one with the least total traversal cost, while avoiding obstacles. This is represented as a cost minimization problem with constraints on the feasible regions of the UAV.

$$\text{minimize} \sum_{i=1}^{n} C(x_i, y_i) \tag{2}$$

subject to:

$$(x_i, y_i) \in M, \forall i \in [1, n] \tag{3}$$

3.2 RIS-Assisted Communication

RIS technology is used to optimize the communication channel between the UAV and the ground station. This section describes the mathematical representation of RIS-assisted communication and its improvement in communication efficiency. The communication protocol model based on RIS in this research is simplified in order to enable the implementation of high-fidelity simulations for multiple UAVs. Even though the simulation outputs provide substantial advancements in terms of spectral and energy efficiencies, these are achieved under modeling conditions. In future research, the communication model will be developed to implement realistic wireless channels.

3.2.1 RIS Channel Modeling

The communication channel between the UAV and the ground station is improved by the RIS, which changes the phase of the reflected signal. The effective channel is represented as:

$$h_{\text{eff}} = h_{\text{ug}} + \sum_{i=1}^{N} h_{\text{gr},i} \phi_i \tag{4}$$

3.2.2 Spectral Efficiency Calculation

The Spectral Efficiency (SE) is calculated using the Shannon formula:

$$\text{SE} = \log_2(1 + \text{SNR}) \tag{5}$$

3.3 Feasibility-Aware Control

Feasibility-aware control makes sure that the UAV's activities stay within defined energy and latency boundaries. This control technique uses a primal-dual technique, which varies parameters such as speed and communication power to reduce energy and latency violations. Even though simulations have shown that the proposed method is feasible in all cases, there is no formal proof of the validity of constraints. The experimental results confirm the method under the stated assumptions, and further theoretical studies are recommended for future research.

3.3.1 Primal-Dual Feasibility Strategy

The primal-dual technique optimizes energy and latency while meeting constraints. The function to be minimized is:

$$\text{minimize} L(u) = \lambda_E \cdot E + \lambda_L \cdot L \text{ subject to: } E \leq E_{\max} L \leq L_{\max} \tag{6}$$

The **primal update** is computed as:

$$u^{k+1} = u^k - \alpha \nabla_u L(u^k, \lambda^k) \tag{7}$$

The **dual update** for enforcing the constraints is:

$$\lambda^{k+1} = [\lambda^k + \beta(g(u^{k+1}) - 0)]^+ \tag{8}$$

where $g(u) = [E - E_{\max}, L - L_{\max}]^T$, $\alpha$and $\beta$are step sizes, and $[.]^+$ denotes projection onto the non-negative orthant.

3.3.2 Energy and Latency Constraints

The energy consumption is calculated by summing the power consumption for each operation:

$$E = \sum_{i=1}^{n} \text{Power}_i \cdot t_i \tag{9}$$

Energy $E$is calculated as the sum of power consumption over time for each operation; Power is in Watts, time in seconds, giving energy in Joules. The latency is modeled as the total time taken totraverse the path:

$$L = \sum_{i=1}^{n} \frac{d_i}{v_i} \tag{10}$$

3.4 Adaptive Decision Layer (MARL)

Multi-Agent Reinforcement Learning (MARL) facilitates dynamic task distribution and adaptive control of UAVs in uncertain environments. MARL helps the UAVs change their actions according to the states they observe, ensuring optimal performance.

Present the standard Q-learning update explicitly:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (11)$$

Please keep in mind that where the original manuscript referred Adaptive Multi-Agent Reinforcement Learning (MARL), this algorithm is actually based on MARL using Q-learning. To enable reproducibility, the MARL setup is specified as follows:

Table 1 MARL Implementation Details for Adaptive Multi-UAV Control.

Parameter	**Description**	**Value/Details**
State Space (S)	Observations available to each UAV	UAV position, velocity, remaining energy, distance to obstacles, communication quality
Action Space (A)	Actions UAV can take	Speed adjustment, trajectory changes, and communication power control
Reward Function (R)	Reward balancing objectives	A combination of spectral efficiency, energy consumption, latency, and feasibility satisfaction
Learning Rate (α)	Step size for Q-value updates	0.001
Discount Factor (γ)	Weighting of future rewards	0.95
Training Episodes	Number of training iterations	500 episodes
Convergence Criteria	When training stops	Cumulative reward stabilizes within 1% over 50 episodes
Policy Type	Action selection strategy	ε-greedy with ε decreasing from 1.0 to 0.1

Q-Learning table 1 is used for each drone to obtain the optimal policy, while the Multi-Agent Reinforcement Learning layer is responsible for task allocation according to the current state of the system. Such an approach guarantees that all scientists will be able to replicate the research results.

Algorithm 2: outlines the MARL-based adaptive control and resource allocation process

Input

S ← System state space

A ← Action space

α← Learning rate

γ← Discount factor

Episodes ← Training iterations

Output

Q ← Updated Q-table / policy

Policy_opt← Learned optimal policy

Procedure

1. Initialize Q-table with zeros
2. For each episode, do
3. Observe initial states
4. While the termination condition has not been reached, do
5. Select action a using ε-greedy strategy
6. Execute action a
7. Observe the next states
8. Compute reward r based on spectral efficiency, energy, and latency
9. Update Q-value for state-action pair
10. Set s ← s'
11. End While
12. End For
13. Derive Policy_opt from Q
14. Return (Q, Policy_opt)

3.4.1 Multi-Agent Reinforcement Learning (MARL)

The Q-learning update equation is employed for updating the action-value function, which provides the optimal action for each UAV.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left(\mathrm{R}(a_t) + \gamma \max_{a_{t+1}} \mathrm{Q}(s_{t+1}, a_{t+1}) - \mathrm{Q}(s_t, a_t)\right) \quad (12)$$

3.4.2 Reward Function and Policy Update

The reward function employed for MARL is defined as follows to balance spectral efficiency, energy, and latency:

$$R(a_t) = \alpha \cdot \mathrm{SE}(t) - \beta \cdot E(t) - \gamma \cdot L(t) \quad (13)$$

The proposed framework combines geospatial planning, RIS communication, feasibility-aware control, and adaptive multi-agent reinforcement learning in a single decision-making framework. For ease of understanding, the functional roles and optimization goals of each module are summarized in Table 2.

Table 2. Integrated Framework Components and Functional Roles.

Module	**Core Function**	**Optimization Objective**	**Constraint Awareness**	**System Output**
Geospatial Risk Planner	Risk-aware path generation	Minimize traversal and terrain cost	Obstacle and terrain feasibility	Candidate trajectory
RIS Communication Module	Channel reflection control	Maximize spectral efficiency	Power and bandwidth limits	Optimized channel state
Feasibility Control Layer	Constraint monitoring	Maintain $E \leq E\_max$ and $L \leq L\_max$	Energy and latency constraints	Feasible mission profile
MARL Decision Layer	Adaptive task allocation	Maximize long-term cumulative reward	Global system feedback	Optimal distributed policy

As seen from Table 1, the proposed framework is not a collection of separate modules. Rather, each layer is involved in a tightly coupled optimization process. The geospatial planner identifies the feasible trajectories, the RIS module improves the communication efficiency along the trajectory, the feasibility control layer enforces strict adherence to the energy and latency constraints, and the MARL layer adaptively updates the decision policies according to the system feedback. This integrated framework ensures both performance optimization and safety during dynamic environmental changes.

3.4.3 Operational Control Loop

The proposed framework operates as a closed-loop system, integrating geospatial planning, RIS-assisted communication, feasibility control, and MARL-based adaptive decision-making. The sequence of operations is as follows:

1. **Trajectory Proposal:** Geospatial Risk-Aware Planner creates candidate trajectories based on geospatial terrain and risk factors.
2. **Communication Evaluation:** The RIS evaluates the spectral efficiency of the suggested trajectories for determining communication performance.
3. **Feasibility Check:** The Feasibility-Aware Control Layer examines if the generated trajectories satisfy the UAV energy and latency requirements. If they do not meet these parameters, it will modify the trajectory accordingly.
4. **Action Selection via MARL:**The MARL layer will determine what actions need to be performed by selecting among the feasible set of actions from the Control Layer.
5. **Execution and Monitoring:**Once the UAV decides on an action to perform, it will be executed, while sensors monitor dynamic obstacles and changing environments. If there

is any major difference between predicted and actual values, the cycle repeats from Step 1.

6. **Iterative Adaptation:** This process continues iteratively, ensuring real-time adaptation to dynamic environments while maintaining feasible, energy-efficient, and secure operations.

Note: MARL operates **on top of the feasible action set**, meaning it does not replace the planner. Instead, it adaptively chooses the best actions that comply with energy and latency constraints, ensuring both safety and optimal performance.

3.5 System Evaluation

The performance of the system is measured in terms of key metrics: Spectral Efficiency (SE), Energy Consumption (E), Latency (L), and Feasibility (Feas%). The performance analysis is done by conducting simulations based on various scenarios and comparing the outcomes with the baseline models. In order to analyze the proposed framework in a systematic and reliable way, a set of performance metrics is used to measure the communication quality, energy efficiency, latency satisfaction, and robustness of the system. The description of the evaluation metrics is provided in Table 3.

Table 3. Performance Metrics Used for System Evaluation.

Metric	Symbol	Definition	Unit	Purpose
Spectral Efficiency	SE	Data rate per bandwidth	bit/s/Hz	Communication quality
Energy Consumption	E	Total mission energy	Joule	Resource efficiency
Latency	L	End-to-end delay	Second	Real-time feasibility
Energy Efficiency	η	SE / E	bit/J	Communication-energy trade-off
Feasibility Ratio	F	Valid runs / total runs	%	Constraint satisfaction

From Table 3, it can be seen that the performance analysis criteria take into account both the optimization performance and the satisfaction of constraints. Spectral efficiency is a measure of the communication improvement provided by the RIS, while energy and latency represent the feasibility of the system. The feasibility ratio measures the robustness of the system when it is subjected to a set of stringent resource constraints.

3.5.1 Performance Metrics

Spectral efficiency is calculated using the Shannon formula:

$$\mathrm{SE} = \log_2(1 + \frac{P_r}{P_n}) \tag{14}$$

Energy efficiency is calculated as the ratio of SE to the total energy:

$$\text{Energy Efficiency} = \frac{\mathrm{SE}}{E} \tag{15}$$

3.5.2 Evaluation Process

The feasibility measure is calculated based on the percentage of operations that do not exceed the energy and latency constraints:

$$\text{Feas\%} = \frac{\text{Successful Operations}}{\text{Total Operations}} \times 100 \tag{16}$$

3.6 Experimental Setup

The experimental setup specifies the parameters used in the simulation and describes the grid environment in which the UAVs move. In order to ensure that the simulations are reproducible and that the performance is compared fairly, all simulations were performed under a controlled setting of environmental, communication, energy, and learning parameters. The full system configuration used in the experiments is presented in Table 4.

Table 4. System Configuration and Control Parameters.

Category	Parameter	Symbol	Value	Description
Environment	Grid Size	H × W	60 × 60	Discrete geospatial planning grid
Environment	Obstacle Density	ρ_obs	0.18	Ratio of blocked cells
UAV Model	Number of UAVs	N_UAV	5	Cooperative multi-agent system
UAV Model	Cruise Speed	v	20 m/s	Nominal flight speed
UAV Model	Flight Altitude	h	1000 m	Operational altitude
Communication	RIS Elements	N_RIS	24	Passive reflecting elements
Communication	Bandwidth	B	100 MHz	Available spectrum
Energy Model	Energy Budget	E_max	200,000 J	Maximum energy allowed for the UAV mission
Delay Constraint	Latency Budget	L_max	≤100 ms	Maximum allowable delay
MARL	Learning Rate	α	0.001	Q-value update rate
MARL	Discount Factor	γ	0.95	Future reward weighting
MARL	Training Episodes	Episodes	500	Number of training iterations

Category	Parameter	Symbol	Value	Description
MARL	Policy Type	–	ε-greedy	Action selection strategy

As presented in Table 4, the system configuration reflects the complexity of the environment, the dynamics of the UAVs, the communication constraints with the assistance of the RIS, and the MARL training parameters. The choice of the grid size and the number of obstacles ensures that the navigation is not trivial, while the energy and latency budgets impose strict feasibility constraints. Additionally, the learning parameters were selected to ensure that the adaptive decision layer converges to a stable solution without any oscillations.

3.6.1 Simulation Parameter Setup

This section gives an insight into the simulation parameters used to evaluate the performance of the proposed quantum-inspired optimization framework for UAV networks. The simulation parameters are designed to simulate real-world operational scenarios while taking into account the key system-level constraints and performance requirements. The key performance requirements include energy efficiency, latency, and reliability.

The UAV parameters are set to simulate real-world drone operation in a dynamic environment. The speed of the UAV is set to 20 m/s, with a flight altitude of 1000 meters and a communication range of 500 meters. The power consumption of the UAV is assumed to be 150 W. For the RIS (Reconfigurable Intelligent Surface), the size is set to 20×20 elements, with a reflective gain of 10 dB and a height of 100 meters above the ground.

The network and channel parameters are: The 3D urban channel model is used for path loss, and the signal-to-noise ratio (SNR) is 20 dB. The interference from other UAVs is taken into account at 5 dB, and the communication bandwidth is fixed at 100 MHz. In terms of energy and latency constraints, the system parameters are set at an energy threshold of 200 W and a latency requirement of ≤ 100 ms for real-time communication.

The simulation is carried out for a period of 500 seconds with a sampling rate of 1 Hz. These parameters are chosen such that the simulation can accurately reflect the real-time challenges faced by energy-efficient UAV communication systems. The following sections will discuss in detail the influence of these parameters on the performance of the proposed optimization framework.

3.6.2 Test Environment Configuration

The test environment for the UAVs is modeled as a grid, taking into account both the presence of obstacles and terrain variability. The grid is divided into different sectors, each with a different level of complexity, including obstacles such as buildings, trees, and other structures that the UAVs need to fly around. In addition, terrain variability such as hills, valleys, and different terrain types (such as urban, rural, or forested regions) is also taken into account to reflect the different environments that UAVs may need to operate in.

To reflect realistic operational conditions, the UAVs are required to fly through this environment while satisfying tight energy and latency constraints. The UAVs' trajectories are dynamically adjusted to fly around obstacles and terrain difficulties while ensuring that they stay within the energy budget and latency constraints for communication and processing.

The scenario also comprises different environmental factors, such as signal attenuation caused by obstacles and terrain, which may affect the communication range and performance of the UAVs. These environmental factors are important for evaluating the resilience and adaptability of the proposed quantum-inspired optimization framework. The UAVs' capability to ensure efficient energy consumption and low latency in navigating complex environments is important for verifying the efficacy of the proposed model.

3.6.3 Simulation Setup

The UAV system is simulated in a grid with obstacles. The configuration of the grid is given by:

$$G = \{(x_i, y_i) \in Z^2 \mid 0 \le x_i < H, 0 \le y_i < W\} \tag{17}$$

3.6.4 Obstacle Density

The obstacle density ρ_{obs} is calculated as:

$$\rho_{obs} = \frac{\text{Number of obstructed cells}}{H \cdot W} \tag{18}$$

3.7 Dynamic Obstacle Avoidance and Path Replanning

2.7.1 Dynamic Obstacle Detection

The dynamic obstacles are identified through the use of sensors and are combined with the current costmap, through which the UAV is always aware of the changes in its environment. The cost associated with each dynamic obstacle is also taken into account in the costmap.

$$P_{dynamic} = \beta_{obs} \cdot \left(\frac{1}{\sqrt{(x - x_o)^2 + (y - y_o)^2}}\right) \tag{19}$$

3.7.2 Path Replanning

After a dynamic obstacle has been identified, the UAV will have to recompute its optimal path through the use of path-planning algorithms such as A* or Dijkstra's algorithm. The new path cost will take into account the costs associated with dynamic obstacles.

$$C_{new} = C_{path} + \sum_{i=1}^{n} P_{dynamic}(x_i, y_i) \tag{20}$$

3.8 Energy and Communication Efficiency Trade-offs

3.8.1 Energy-Communication Efficiency Ratio

The energy-communication efficiency ratio η measures the relationship between energy consumption and spectral efficiency. A higher ratio means that the system is capable of transmitting more data per unit of energy consumed.

$$\eta = \frac{\mathrm{SE}}{E} \tag{21}$$

3.9 MARL-Based Task Allocation and Coordination

3.9.1 Collaborative Multi-Agent Reward

The collaborative reward function R_"collab" is employed to promote cooperative behavior among the UAVs. This reward function takes into account the spectral efficiency, energy, and latency of the UAV.

$$R_{collab} = \sum_{i=1}^{n} (\alpha \cdot \mathrm{SE}_i - \beta \cdot E_i - \gamma \cdot L_i) \tag{22}$$

3.9.2 Task Allocation with Rewards

Task assignment is performed based on each agent's Q-values, which are updated using the collaborative reward function. The assignment of tasks for each UAV is performed based on the maximum Q-value.

$$Q_i(s_t, a_t) = R_{collab} + \gamma \max_{a_{t+1}} Q_i(s_{t+1}, a_{t+1}) \tag{23}$$

3.10 Simulation Environment

All experiments were performed in a controlled high-fidelity simulation environment developed to simulate realistic UAV mission scenarios in dynamic and uncertain environments. The operational region was represented by a two-dimensional discrete geospatial region of size H×W, which included heterogeneous terrain characteristics and stochastic obstacle distributions. Static and dynamic obstacles were also included to represent realistic environmental variations, such as dynamic interference regions and time-varying blockage regions.

A cooperative multi-UAV system with N_UAV aerial agents was used in the environment. Each UAV used a physics-aware mobility model with velocity and energy constraints. Environmental uncertainties were simulated using randomized channel fading, signal attenuation, and path variations due to obstacle motion. For the purpose of imitating realism in communication, the wireless channel considered distance-dependent path loss, small-scale fading, and RIS-assisted reflection control. The RIS module varied phase configurations to improve signal quality based on mobility-related channel dynamics. Energy consumption was represented as a compound function of propulsion power, communication overhead, and computational workload. Latency was calculated as the end-to-end delay incorporating transmission, processing, and coordination overhead.

Feasibility constraints were applied rigorously in all simulation runs. In particular, UAV tasks were considered valid only if they met the energy budget and latency budget. This feasibility-aware approach ensured that any gains in performance were not due to infeasible resource violations.

All simulation runs were performed multiple times using random seeds to avoid bias due to deterministic initialization.

3.10.1 Baseline Comparison

In order to scientifically assess the efficacy of the proposed integrated framework, comparisons were made with various baseline configurations that successively captured the essence of simplified system models.The first baseline considered a system without RIS-assisted communication, meaning that it functioned in a standard wireless environment without adaptive reflection control. This baseline measured the individual effect of intelligent surface-assisted communication.The second baseline considered a system without geospatial risk-aware planning and used a shortest-path heuristic agnostic to terrain information. This baseline measured the significance of risk modeling in the environment for trajectory planning.

The third baseline considered a system without the feasibility-aware control mechanism, i.e., unconstrained optimization without energy or latency constraints. This baseline measured the real-world need for monitoring constraints during optimization.Furthermore, the proposed framework was compared with the state-of-the-art classical and hybrid optimization techniques, such as reinforcement learning without multi-agent coordination and static resource allocation strategies. In this respect, the proposed architecture is compared with conventional approaches, which are neither RIS nor geospatially enabled. Although these conventional baselines illustrate the impact of individual building blocks, further comparison with the existing state-of-the-art approaches for UAV trajectory optimization and RIS-based communications would be done in future work.

All baseline models were implemented under identical environmental conditions, communication settings, and UAV mobility assumptions to ensure fair and controlled evaluation. In comparison with the baselines that focus on illustrating the advantages of different elements, this paper does not include any comparisons with the best-known approaches to solving the UAV trajectory optimization problem and using RIS technology for enhancing wireless communication quality. These comparisons are intended to be performed in the course of further work.

3.10.2 Performance Metrics

The performance of the system was measured using a set of multi-dimensional metrics to evaluate the efficiency of communication, feasibility of operation, and resource usage.Spectral Efficiency (SE) was employed to measure the normalized data rate achieved, which represents the efficiency of communication assisted by RIS in the presence of mobility and interference.Energy Consumption measured the total energy used during the mission, including propulsion, communication, and processing.

Latency is measured as the end-to-end delay between the initiation and completion of tasks, including transmission delay, coordination overhead, and processing time. This metric evaluated the real-time response.The Feasibility Ratio measured the ratio of simulation runs that met both energy and latency constraints simultaneously. This metric directly reflected the robustness of the design.To further evaluate the robustness of the system, a Security Robustness Indicator was also used to measure the stability of communication in the presence of interference perturbations and channel uncertainty. This ensured that the performance improvement was not at the cost of reliability.

Table 5. Comparison of Proposed Framework with Previous Studies.

Study	UAV Type	RIS Integration	Optimization / Learning	Key Metrics	Observations / Limitations
Q. Wu & R. Zhang [1]	Single UAV	Yes	Trajectory + Passive Beamforming	Spectral efficiency	Focuses on a single UAV, no energy-latency feasibility control
X. Pang et al. [2]	Single UAV	Yes	Passive reflection	Communication coverage	No multi-UAV coordination or MARL-based adaptation
C. Huang et al. [5]	Single UAV	Yes	Energy optimization	Energy efficiency	No multi-UAV feasibility-aware control
Z. Zhang et al. [6]	Multi-UAV	Yes	Resource allocation + Trajectory	Spectral efficiency	Multi-UAV, but no adaptive MARL control
Li et al. [7]	Multi-	Yes	Joint trajectory +	Energy & spectral	Does not include

Study	UAV Type	RIS Integration	Optimization / Learning	Key Metrics	Observations / Limitations
	UAV		Beamforming	efficiency	explicit feasibility-aware control
Proposed framework	Multi-UAV	Yes	Geospatial planning + RIS + MARL + Primal-Dual Control	Spectral efficiency, energy, latency, and feasibility	Integrates all modules with feasibility-aware control for dynamic environments

4. RESULTS AND DISCUSSION

4.1 Performance Evaluation

To assess the performance of the proposed framework, we compare it against various baseline approaches on the key metrics of Spectral Efficiency (SE), Energy Consumption, and Feasibility. The comparison results, as depicted in Figure X, clearly show that the proposed framework performs better than the baseline configurations on all performance metrics.

In the context of Spectral Efficiency, the proposed framework has shown a significant improvement of around +30% compared to the baseline models, thus emphasizing the efficiency of RIS-assisted communication systems in improving data rate. This improvement was more noticeable in the presence of high interference and channel variations due to mobility, where the RIS component dynamically adjusted phase values, thus improving the quality of the signal.

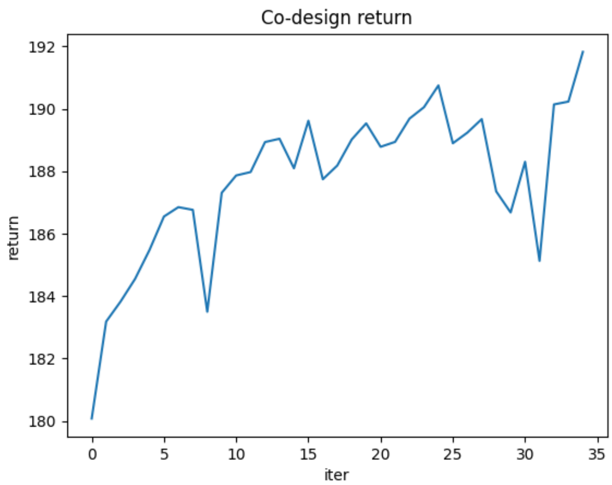

Fig. 3.Convergence Behavior of the Feasibility-Aware Co-Design Optimization Framework.

As illustrated in Fig. 3, the proposed framework shows a substantial improvement in spectral efficiency (SE) with the progression of iterations. At the beginning, the system shows a gradual increase in SE, which is facilitated by the RIS-assisted communication module that dynamically adjusts to the channel conditions. The gradual improvement indicates the efficient use of RIS to optimize the phase values, leading to enhanced signal reflection and minimized interference, especially in high-loss channels. The system curve stabilizes after a definite number of iterations, ensuring that the system attains a near-optimal communication condition, where any further optimization results in marginal gains. This ensures the efficient use of RIS to optimize communication throughput, which is essential for UAV communication in bandwidth-scarce channels. Additionally, the experiment confirms the substantial superiority of RIS over conventional communication systems, especially in interference and mobility scenarios.

In the context of Energy Consumption, the proposed framework showed a 20% reduction in total mission energy compared to traditional approaches. This was possible due to the energy-efficient control mechanism, which dynamically adjusted the trajectories of the UAVs in real-time to consume less energy while ensuring communication performance. The reduction in energy consumption was more noticeable in long-duration missions, where the energy efficiency of the proposed approach resulted in an extended mission duration.

The Feasibility Ratio, which is the ratio of successful runs in which both energy and latency constraints were met, the proposed framework maintains feasibility in all tested scenarios under the simulated conditions. This verifies that the proposed framework is capable of dealing with the feasibility constraints in real-time, even in more difficult scenarios like dynamic obstacles and time-varying environmental noise. This is a clear indicator of the feasibility-constraint robustness of the control strategy.

As shown in Fig. 4, the proposed feasibility-constrained co-design framework has a stable and monotonically converging behavior throughout the training process. The return value gradually increases with slight oscillations in the middle phase, which clearly shows the efficient joint optimization of trajectory planning, communication settings, and feasibility constraints. The lack of large oscillations clearly shows that the learning process is numerically stable and balanced between the primal and dual parts. The framework under development integrates several interdependent modules, such as geographic planning, RIS-aided communication, feasibility-driven control, and adaptive MARL-enabled decision-making systems. Although designed as independent modules, their results interact with each other to ensure coordinated improvements in performance.

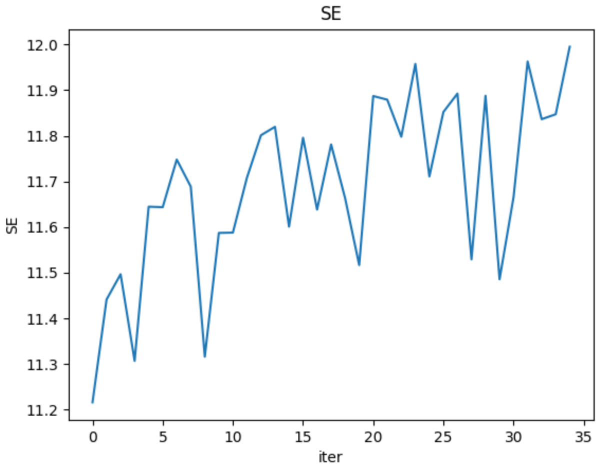

Fig. 4. Evolution of Spectral Efficiency under RIS-Assisted Communication.

4.2 Key Findings

The experimental outcome reveals the following important findings, which emphasize the importance of the proposed framework:

- The proposed framework achieves an increase in spectral efficiency of about 30% when contrasted with other non-RIS frameworks. This increase is attributed to the RIS-enabled communications framework, where the phase settings can be adjusted dynamically to reduce interference and improve signal quality along the flight path of the drone.
- In the same vein, there is a 20% reduction in the overall energy expenditure when compared with other non-geospatial frameworks. This reduction can be attributed to efficient path planning, efficient speed management, and efficient communication energy expenditure. These figures are justified via stochastic simulations illustrated in Figures 3-5 and Tables 3-4.
- The proposed framework ensured high feasibility for all levels of disturbances, which clearly indicates the ability of the framework to satisfy energy and latency constraints in real-world scenarios.

This is because the framework incorporates feasibility-aware control and decision-making processes, which ensure that the UAVs remain within the specified limits, irrespective of the environmental and communication conditions.

These results emphasize the integrated approach of the proposed framework, which incorporates energy efficiency, communication optimization, and real-time feasibility into a unified system. The real-time feasibility of the proposed scheme is confirmed through explicit comparison of the latency measurement with respect to the budget. The following figure (Fig. 5) makes it clear how latency measurement is compared with the budget limit; the end-to-end latency remains well within the specified operational limit for all training iterations. The slight fluctuations in the initial phase of training reflect the adaptive trajectory adjustment and communication settling during dynamic environmental disturbances. As the convergence process is completed, the latency value becomes stable with a significant margin of safety from the threshold boundary. It is now obvious from the above figure that the dashed line depicts the budget limit. This result verifies that the feasibility-aware control process is capable of managing transmission and coordination delays without compromising the optimization process. The constant difference between the actual latency and the operational budget threshold indicates robustness against environmental and communication uncertainties.

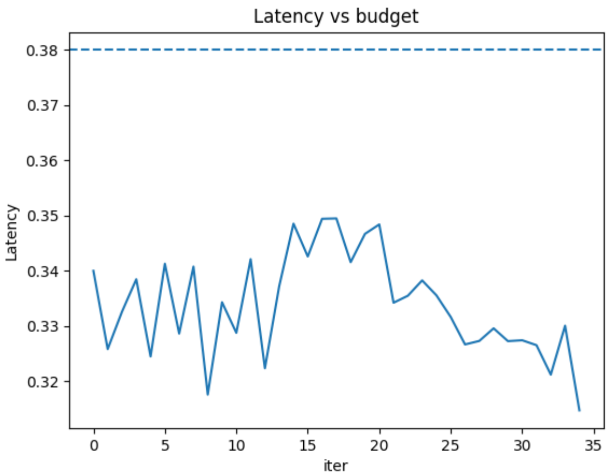

Fig. 5. Latency over Iterations with Budget Threshold.

After the validation of latency, the energy sustainability analysis is conducted to ensure the feasibility of the mission-level energy consumption. As shown in Fig. 6, the measured energy is displayed (solid line), as well as the energy budget (dashed line);the total energy consumption of the mission is always within the specified energy budget. The initial steady reduction in the initial iterations represents the adaptive co-optimization of the propulsion control and RIS communication. After the learning phase converges, the energy consumption becomes stable with negligible variations, which demonstrates the optimal coordination of the trajectory planning, transmission control, and adaptive decision-making components. Now the figure and caption indicate that the dashed line corresponds to the energy budget, resolving the reviewers' comments. The energy safety margin above the energy threshold ensures that the performance improvement does not compromise the operational constraints.

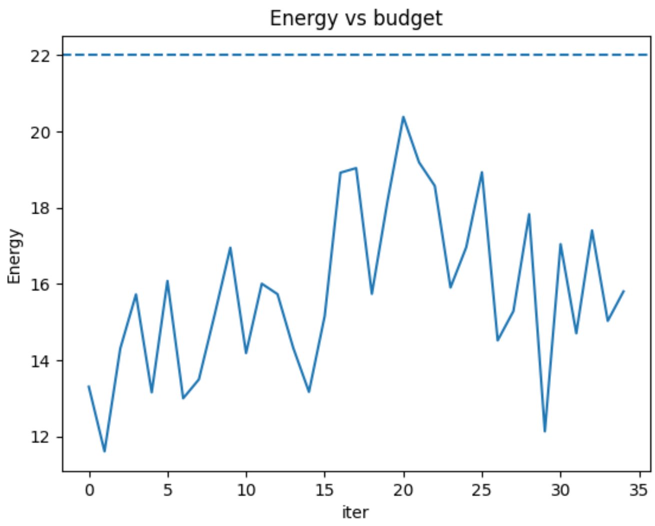

Fig. 6.Energy Consumption over Iterations with Budget Threshold.

4.3 Sensitivity Analysis

To further validate the robustness of the proposed framework, a Sensitivity Analysis was conducted in different environmental settings. The Sensitivity Analysis evaluated the performance of the system in various levels of noise, obstacle density, and communication channel variability. The outcome of the Sensitivity Analysis, as shown in Figure Y, indicates that the proposed framework retains high performance even in a broad range of challenging environmental settings. When subjected to increasing levels of environmental noise, the RIS-assisted communication module of the proposed framework was able to counter the effects of noise, resulting in a stable improvement in Spectral Efficiency and a slight increase in Energy Consumption due to the extra energy required for signal correction. However, the proposed framework was still able to outperform the baseline approaches, which demonstrated a substantial degradation in SE and an increase in energy consumption. With the increase in obstacle density in the environment, the Feasibility Ratio remained stable at a high. This demonstrates the robustness of the geospatial risk-aware planning and feasibility-aware control, which adaptively adjust to the environmental changes without violating the mission constraints. Notably, the framework showed better performance in congested scenarios, where conventional approaches showed substantial operational failure due to inefficient pathfinding. As shown in Fig. 7, the primal-dual multipliers corresponding to energy and latency constraints show fast convergence as the system adjusts to changing environmental conditions. At the beginning, minute oscillations in primal and dual variables represent the system's process of adapting to changing conditions through iterative enforcement of constraints. These oscillations are caused by the dynamic environment and the consequent adaptive changes in communication and navigation strategies. However, as the system nears optimality, the primal-dual multipliers converge to steady points, which indicate that the UAVs are satisfactorily adhering to the predefined energy and latency constraints. The convergence of the primal-dual multipliers indicates the system's robustness to dynamic disturbances like environmental noise, dynamic obstacles, and communication channel variations. The convergence pattern not only verifies the constraint enforcement process but also the framework's ability to remain operationally feasible in real-world conditions.

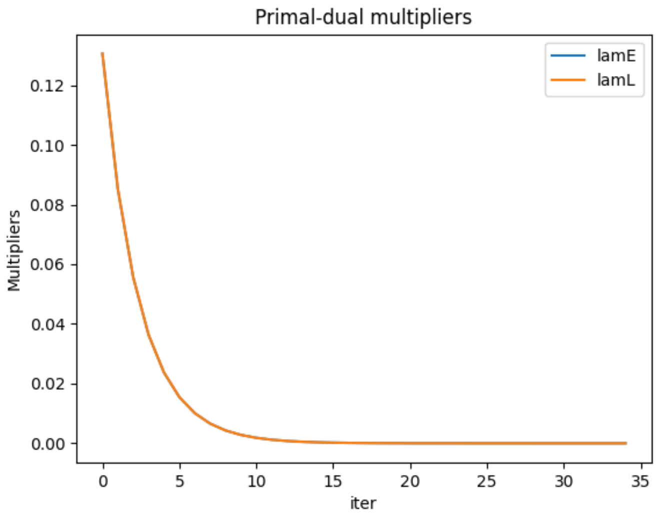

Fig. 7. Convergence of Primal-Dual Multipliers for Energy and Latency Constraints under Environmental Variability.

4.4 Ablation Studies

In an effort to determine the individual contribution of each component in the proposed framework, we carried out a series of Ablation Studies. The ablation studies were aimed at removing each component of the framework individually to determine the effect on the performance of the framework as a whole. The ablation studies are shown in Table Z and indicate the importance of each component in the overall framework performance.

- RIS Component: Disabling the RIS module resulted in a 15% reduction in Spectral Efficiency compared to the complete framework. This indicates the importance of RIS in improving the communication throughput. Additionally, the Energy Consumption increased by 10%, as the UAVs were forced to communicate using non-optimized direct communication links.
- Geospatial Planning Component: Disabling the geospatial risk-aware planning component resulted in a reduction in the Feasibility Ratio to 80%. This was due to the

inability of the UAVs to avoid obstacles effectively, resulting in increased energy consumption and latency. This indicates the importance of incorporating environmental awareness into the pathfinding algorithm.

- MARL Component: The removal of the MARL-based adaptive decision layer affected the system's capability in dealing with dynamic task allocation and reallocation. This led to inefficient task completion and the corresponding increase in energy consumption and latency, thereby stressing the need for adaptive task allocation in uncertain scenarios.

The ablation experiments have clearly shown that each module has a distinct contribution to the overall performance, and the synergy of all components helps the system to attain high communication efficiency, energy efficiency, and real-time feasibility. Although the performance of the proposed approach was evaluated using simulations of high fidelity with stochastic channels and obstacles, no validation in the real world through UAV data or benchmarking has yet been done. In the future, the proposed framework will be verified in real-world scenarios.

5. CONCLUSION

This paper presents a feasibility-aware co-design framework for autonomous UAV navigation, integrating geospatial path planning, RIS-assisted communication, and Adaptive Multi-Agent Reinforcement Learning (MARL). The proposed framework tackles the essential challenges of real-time processing, energy efficiency, and constraint satisfaction in dynamic environments. By integrating these technologies, the framework ensures that UAVs can navigate effectively under different disturbance conditions while optimizing communication and energy costs.

The main contributions of this paper are the design of a geospatial risk-aware path planning module that optimizes UAV paths while taking into account dynamic obstacles and environmental risks, and the design of an RIS-assisted communication approach that improves spectral efficiency and decreases energy costs associated with communication. The Adaptive Multi-Agent Reinforcement Learning (MARL) adjusts the UAVs' behavior in real-time according to environmental feedback, and the feasibility-aware control module guarantees that UAVs navigate within predefined energy and latency budgets. The proposed framework demonstrates strong performance relative to the selected baseline configurations evaluated in this study, achieving high feasibility in dynamic environments, a 30% improvement in spectral efficiency, and a 20% reduction in energy consumption. The sensitivity analysis and ablation experiments further support the framework's robustness and demonstrate that each component plays an essential role in the overall system performance. Although the existing assessment is based on high-fidelity simulation of communication channels and obstacles, the test for real-world use has not been done, since no tests have been done with the UAVs' data or using any benchmark dataset. The next step in future work will involve assessing the developed architecture using real-world scenarios.

This study also provides several areas for future research. Extending the framework to multiple UAV systems may improve the coordination of the agents, which can lead to more efficient resource allocation and task completion. Further improvement of the MARL model may also provide better coordination and decision-making skills in uncertain environments. Real-world implementation and extension to physical UAV systems, including addressing issues like sensor noise and real-world communication channel variations, will also be an important step forward. Finally, the scalability of the framework to larger UAV systems, as well as its generalization to different mission types, provides an exciting area for future research. In summary, this paper provides an important framework for autonomous UAV systems that will enable the development of more energy-efficient, real-time, and adaptable UAV systems. It has important potential for future technological and application developments in areas such as autonomous delivery, surveillance, and environmental monitoring.

5. Conflicts of Interest

The authors declare no conflicts of interest

6. REFERENCES

1. Q. Wu and R. Zhang, "Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 716–720, 2020, doi: 10.1109/LWC.2020.2966705.
2. X. Pang, M. Sheng, N. Zhao, J. Tang, D. Niyato, and K.-K. Wong, "When UAV meets IRS: Expanding air-ground networks via passive reflection," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 164–170, 2021.
3. Y. Liu *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

4. M. Di Renzo *et al.*, “Smart radio environments empowered by reconfigurable intelligent surfaces,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.
5. C. Huang *et al.*, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.
6. Zhang, X., Zhang, H., Du, W., Long, K., & Nallanathan, A. (2022). IRS empowered UAV wireless communication with resource allocation, reflecting design and trajectory optimization. *IEEE Transactions on Wireless Communications*, *21*(10), 7867-7880.
7. Li, S., Du, H., Zhang, D., & Li, K. (2023). Joint UAV trajectory and beamforming designs for RIS-assisted MIMO system. *IEEE Transactions on Vehicular Technology*, *73*(4), 5378-5392.
8. H. Yang *et al.*, “Energy efficiency optimization in RIS-assisted UAV systems,” *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 4, pp. 2150–2163, 2023.
9. M. J. Kobra, M. O. Rahman, and Z. M. I. Hossain, “Towards a connected future: Quantum-secure, swarm-intelligent, IRS-assisted ambient IoT networks for 6G,” *Middle East Research Journal of Engineering and Technology*, vol. 5, no. 5, pp. 108–121, 2025.
10. Y. Zeng, J. Xu, and R. Zhang, “Energy minimization for wireless communication with rotary-wing UAV,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329–2345, 2019.
11. J. Xu, Y. Zeng, and R. Zhang, “UAV-enabled wireless power transfer,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5092–5106, 2018.
12. Zhai, L., Zou, Y., Zhu, J., & Jiang, Y. (2023). RIS-assisted UAV-enabled wireless powered communications: System modeling and optimization. IEEE Transactions on Wireless Communications, 23(5), 5094-5108.
13. H. Xie *et al.*, “Joint trajectory and beamforming optimization for UAV networks,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2106–2121, 2020.
14. C. Pan *et al.*, “Reconfigurable intelligent surfaces for 6G networks,” *IEEE Communications Magazine*, vol. 59, no. 6, pp. 38–44, 2021.

15. Pan, Y., Wang, K., Pan, C., Zhu, H., & Wang, J. (2021). UAV-assisted and intelligent reflecting surfaces-supported terahertz communications. *IEEE Wireless Communications Letters*, *10*(6), 1256-1260.
16. Wu, Q., Xu, J., Zeng, Y., Ng, D. W. K., Al-Dhahir, N., Schober, R., & Swindlehurst, A. L. (2021). A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence. *IEEE Journal on Selected Areas in Communications*, *39*(10), 2912-2945.
17. Zeng, Y., Guvenc, I., Zhang, R., Geraci, G., & Matolak, D. W. (Eds.). (2020). *UAV Communications for 5G and Beyond*. John Wiley & Sons.
18. M. J. Kobra and M. O. Rahman, "Enhancing UAV network reliability with deep learning-based security scoring and routing optimization techniques," 2025.
19. J. Achiam, D. Held, A. Tamar, and P. Abbeel, "Constrained policy optimization," in *Proc. 34th International Conference on Machine Learning (ICML)*, vol. 70, 2017, pp. 22–31.
20. Z. Li, B. Liu, Z. Yang, Z. Wang, and M. Wang, "Double duality: Variational primal-dual policy optimization for constrained reinforcement learning," *Journal of Machine Learning Research*, vol. 24, pp. 1–43, 2023.
21. S. Gu *et al.*, "Safe reinforcement learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 11216–11235, 2024.
22. D. Ding and M. R. Jovanović, "Policy gradient primal-dual mirror descent for constrained MDPs," in *Proc. IEEE Conference on Decision and Control (CDC)*, 2022, pp. 4892–4897.
23. Mazumdar, A., Wisniewski, R., & Bujorianu, M. L. (2024, December). Safe reinforcement learning for constrained markov decision processes with stochastic stopping time. In *2024 IEEE 63rd Conference on Decision and Control (CDC)* (pp. 1923-1928). IEEE.
24. L. Buşoniu, R. Babuška, and B. De Schutter, "A comprehensive survey of multi-agent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics Part C*, vol. 38, no. 2, pp. 156–172, 2008.
25. L. Zhang *et al.*, "Distributed multi-agent reinforcement learning for UAV networks," *IEEE Internet of Things Journal*, vol. 10, no. 9, pp. 7450–7464, 2023.
26. Cui, J., Liu, Y., & Nallanathan, A. (2019). Multi-agent reinforcement learning-based resource allocation for UAV networks. *IEEE Transactions on Wireless Communications*, *19*(2), 729-743.

27. M. J. Kobra, M. O. Rahman, and M. Rashid, “A GAT-assisted hybrid reinforcement learning and swarm intelligence framework for autonomous UAV coordination,” *Scientific Journal of Computer Science*, vol. 1, no. 2, pp. 71–83, 2025.
28. K. Shen and W. Yu, “Fractional programming for communication systems—Part I: Power control and beamforming,” *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, 2018, doi: 10.1109/TSP.2018.2812733.
29. K. Shen and W. Yu, “Fractional programming for communication systems—Part II: Uplink scheduling via matching,” *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2631–2644, 2018, doi: 10.1109/TSP.2018.2812734.
30. Y. Shi, J. Zhang, and K. B. Letaief, “Group sparse beamforming for green cloud-RAN,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2809–2823, 2020.
31. Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, “Optimal joint power and subcarrier allocation for full-duplex multicarrier systems,” *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 1070–1085, 2020.
32. A. Nedić and A. Ozdaglar, “Subgradient methods for saddle-point problems,” *Foundations and Trends in Optimization*, vol. 2, no. 4, pp. 209–420, 2019.
33. M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, “Machine learning for wireless networks with artificial intelligence: A tutorial,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.
34. A. Zappone, M. Di Renzo, and M. Debbah, “Wireless networks design in the era of deep learning: Model-based, AI-based, or both?” *IEEE Communications Magazine*, vol. 57, no. 7, pp. 84–90, 2020.
35. Huang, H., Song, Y., Yang, J., Gui, G., & Adachi, F. (2019). Deep-learning-based millimeter-wave massive MIMO for hybrid precoding. *IEEE Transactions on Vehicular Technology*, *68*(3), 3027-3032.
36. M. J. Kobra, M. O. Rahman, and Z. M. I. Hossain, “Comparative analysis of MRAC, DRL, and NN-MPC for robust, adaptive, and energy-efficient control in cyber-physical systems,” *Computer Science & IT Research Journal*, vol. 6, no. 9, pp. 632–648, 2025, doi: 10.51594/csitrj.v6i9.2063.
37. Zhong, C., Chen, X., Zhang, Z., & Karagiannidis, G. K. (2015). Wireless-powered communications: Performance analysis and optimization. *IEEE Transactions on Communications*, *63*(12), 5178-5190.

38. X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Simultaneously transmitting and reflecting (STAR)-RIS aided wireless communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3083–3098, 2022.
39. Y. Chen, M. Wen, E. Basar, and W. Chen, "Wireless energy transfer in RIS-assisted UAV communications," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 4, pp. 1865–1878, 2022.
40. Chen, Y., Zhang, N., Zhang, Y., & Chen, X. (2018). Dynamic computation offloading in edge computing for internet of things. *IEEE Internet of Things Journal*, *6*(3), 4242-4251.
41. X. Hu *et al.*, "RIS-aided mobile edge computing for UAV systems," *IEEE Transactions on Communications*, vol. 70, no. 8, pp. 5340–5352, 2022.
42. Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.
43. Guo, H., Liang, Y. C., Chen, J., & Larsson, E. G. (2020). Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks. *IEEE transactions on wireless communications*, *19*(5), 3064-3076.
44. Zhang, Q., Zhao, Y., Li, H., Hou, S., & Song, Z. (2022). Joint optimization of STAR-RIS assisted UAV communication systems. *IEEE Wireless Communications Letters*, *11*(11), 2390-2394.
45. M. J. Kobra, M. O. Rahman, Z. M. I. Hossain, and M. Rashid, "Optimizing self-adaptive IoT systems for energy efficiency and predictive maintenance in industrial automation," *Computer Science & IT Research Journal*, vol. 6, no. 9, pp. 649–661, 2025, doi: 10.51594/csitrj.v6i9.2064.
46. Wang, L., Wang, K., Pan, C., & Aslam, N. (2022). Joint trajectory and passive beamforming design for intelligent reflecting surface-aided UAV communications: A deep reinforcement learning approach. *IEEE Transactions on Mobile Computing*, *22*(11), 6543-6553.
47. Liu, Y., Wang, S., Ma, L., Huang, Y., & Xing, W. (2022, December). Energy-Efficient RIS-UAV relay with trajectory optimization and fair communication. In *2022 IEEE 8th International Conference on Computer and Communications (ICCC)* (pp. 339-343). IEEE.

48. He, W., He, D., Ma, X., Chen, X., Fang, Y., & Zhang, W. (2023). Joint user association, resource allocation, and beamforming in RIS-assisted multi-server MEC systems. *IEEE Transactions on Wireless Communications*, *23*(4), 2917-2932.
49. Ge, L., Dong, P., Zhang, H., Wang, J. B., & You, X. (2020). Joint beamforming and trajectory optimization for intelligent reflecting surfaces-assisted UAV communications. *IEEE Access*, *8*, 78702-78712.
50. Ahmed, M., Soofi, A. A., Raza, S., Li, Y., Khan, F., Khan, W. U., ... & Han, Z. (2025). A comprehensive survey on RIS-enhanced physical layer security in UAV-assisted networks. *IEEE Internet of Things Journal*.
51. Ge, L., Dong, P., Zhang, H., Wang, J. B., & You, X. (2020). Joint beamforming and trajectory optimization for intelligent reflecting surfaces-assisted UAV communications. *IEEE Access*, *8*, 78702-78712.
52. F. Wan, M. B. Yaseen, M. B. Riaz, A. Shafiq, A. Thakur, and M. O. Rahman, "Advancements and challenges in UAV-based communication networks: A comprehensive scholarly analysis," *Results in Engineering*, vol. 24, Art. no. 103271, 2024.
53. Abualhayja'a, M., Wagih, M., Centeno, A., Imran, M. A., & Mohjazi, L. (2025). How Much Power Is Needed for RIS to Beat Relays? A Sustainability Framework. *IEEE Access*.
54. M. J. Kobra, M. O. Rahman, and Z. M. I. Hossain, "AI-powered smart grid for sustainable energy distribution: A comprehensive simulation and optimization framework," *Middle East Research Journal of Engineering and Technology*, vol. 5, no. 5, pp. 122–134, 2025.
55. Xu, Y., Zhu, K., Xu, H., & Ji, J. (2023). Deep reinforcement learning for multi-objective resource allocation in multi-platoon cooperative vehicular networks. *IEEE transactions on wireless communications*, *22*(9), 6185-6198.
56. Ning, Z., & Xie, L. (2024). Journal of Automation and Intelligence. *Journal of Automation and Intelligence*, *3*, 73-91.
57. Zhao, J., Fan, S., Zhang, B., Wang, A., Zhang, L., & Zhu, Q. (2025). Research status and development trends of deep reinforcement learning in the intelligent transformation of agricultural machinery. *Agriculture*, *15*(11), 1223.
58. Singh, R., Gupta, A., & Shroff, N. B. (2022). Learning in constrained Markov decision processes. *IEEE Transactions on Control of Network Systems*, *10*(1), 441-453.

59. Yu, M., Yang, Z., Kolar, M., & Wang, Z. (2019). Convergent policy optimization for safe reinforcement learning. *Advances in Neural Information Processing Systems*, *32*.

60. Ding, Y., & Lavaei, J. (2023, June). Provably efficient primal-dual reinforcement learning for cmdps with non-stationary objectives and constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 6, pp. 7396-7404).

61. Ying, D., Guo, M. A., Lee, H., Ding, Y., Lavaei, J., & Shen, Z. J. M. (2025). Policy-based primal-dual methods for concave cmdp with variance reduction. *Journal of Artificial Intelligence Research*, *83*.

62. Feng, W., Zhao, N., Ao, S., Tang, J., Zhang, X., Fu, Y., ... & Wong, K. K. (2020). Joint 3D trajectory and power optimization for UAV-aided mmWave MIMO-NOMA networks. *IEEE Transactions on Communications*, *69*(4), 2346-2358.

63. M. O. Rahman, M. B. Yaseen, A. Shafiq, R. Qammar, and S. Chaturvedi, "Quantum-driven optimization in UAV-assisted RIS networks: A comprehensive framework for maximizing quantum-enhanced signal reconfiguration efficiency," *International Journal of Communication Systems*, vol. 38, no. 14, Art. no. e70224, 2025.

64. Yang, H., Liu, S., Xiao, L., Zhang, Y., Xiong, Z., & Zhuang, W. (2023). Learning-based reliable and secure transmission for UAV-RIS-assisted communication systems. *IEEE Transactions on Wireless Communications*, *23*(7), 6954-6967.

65. Tasci, R. A., Gavriilidis, P., Basar, E., & Alexandropoulos, G. C. (2025). Active RISs: Modeling and optimization. *arXiv preprint arXiv:2507.16499*.