International Core Journal of Engineering Volume 12 Issue 4, 2026
ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

Robust Joint Optimization for Secure Transmission in RIS-Assisted UAV Systems with Attitude Jitter

Xiaojuan Bai*, Xiangtian Liu, Ao Gao, and Xiaoyong Yang

Northwest Normal University, Lanzhou, Gansu, 730070, China

*Corresponding author: Xiaojuan Bai (Email: baixiaojuan@nwnu.edu.cn)

Abstract

This paper investigates secure downlink transmission in RIS-assisted mmWave UAV communication systems under UAV attitude jitter. The attitude jitter is modeled as small-angle rotational perturbations, which lead to array-response mismatch and CSI uncertainty. To address the resulting coupled optimization problem, a dual-agent deep reinforcement learning framework is proposed to jointly optimize UAV trajectory, active beamforming, and RIS phase shifts. Simulation results show that the proposed method improves the average secrecy rate under different jitter conditions and achieves better robustness and convergence performance than benchmark methods. These results demonstrate the effectiveness of the proposed approach for robust secure transmission in RIS–UAV systems.

Keywords

Unmanned Aerial Vehicle; Reconfigurable Intelligent Surface; Physical Layer Security; Attitude Jitter.

1. Introduction

Unmanned Aerial Vehicles (UAVs), with their flexible deployment and high probability of line-of-sight (LoS) links, have shown great potential for future wireless communication systems [1]. However, the broadcast nature of air-to-ground channels introduces serious security risks, especially in LoS-dominant scenarios where eavesdroppers can easily intercept strong signals[2]. Reconfigurable Intelligent Surfaces (RISs), consisting of a large number of low-cost passive elements, have recently attracted significant attention for their ability to manipulate the wireless propagation environment[3]. By adjusting the phase shifts of reflecting elements, RIS can enhance desired signals and suppress interference. Therefore, integrating RIS into UAV systems has become an effective approach to improve both communication performance and physical layer security[4]. Existing works have investigated RIS-assisted UAV systems through joint optimization of UAV trajectory, active beamforming, and RIS phase shifts, typically using iterative optimization methods such as SCA and SDR[5]. In addition, physical layer security techniques, including artificial noise and secure beamforming, have been widely studied to mitigate eavesdropping threats[6]. Recently, deep reinforcement learning (DRL) has been introduced to address the high-dimensional and non-convex nature of these problems, offering a data-driven approach for policy optimization in complex environments[7]. However, most existing studies mainly consider CSI imperfections caused by estimation errors or feedback delay[8], while the impact of UAV attitude jitter has been largely overlooked. In practical scenarios, airflow disturbances and structural vibrations introduce random small-angle rotations of the UAV, leading to deviations in AoA/AoD and mismatches in array responses. This results in inaccurate CSI and performance degradation, which is particularly critical in RIS-assisted mmWave systems. Motivated by this, this paper investigates an RIS-assisted

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

mmWave UAV secure communication system under UAV attitude jitter. The jitter is modeled as small-angle rotational perturbations, which induce CSI uncertainty. To address the resulting coupled optimization problem, a dual-agent DRL framework is proposed to jointly optimize UAV trajectory, beamforming, and RIS phase shifts. Simulation results demonstrate that the proposed method improves secrecy performance and exhibits strong robustness under different jitter conditions.

2. System Model

2.1 Overview of System Models

As shown in Fig. 1, this paper considers a RIS-assisted millimeter-wave UAV secure downlink communication system. In the considered system, the UAV serves K single-antenna legitimate users with the assistance of the RIS, while P single-antenna eavesdroppers may intercept the transmitted signals. The UAV is equipped with a uniform linear array (ULA) with A antennas, and the RIS consists of M passive reflecting elements arranged in a uniform planar array (UPA). The UAV flies at a fixed altitude H_U, and the overall transmission period is divided into N time slots. The UAV position at slot n is denoted by $\mathbf{q}[n]=[x_U[n], y_U[n], H_U]^{\mathrm{T}}$. The RIS location is fixed at $\mathbf{w}_R=(x_R, y_R, z_R)^{\mathrm{T}}$. The positions of legitimate users and eavesdroppers are denoted by $\mathbf{w}_k=[x_k, y_k, z_k]^{\mathrm{T}}$ and $\mathbf{w}_p=[x_p, y_p, z_p]^{\mathrm{T}}$, respectively. The UAV mobility is constrained by its maximum flying distance between consecutive time slots, denoted by $D_{\max}$. In addition, the UAV is restricted to operate within a predefined horizontal region. The initial position of the UAV is set as $\mathbf{q}[0]=\mathbf{q}_0$.

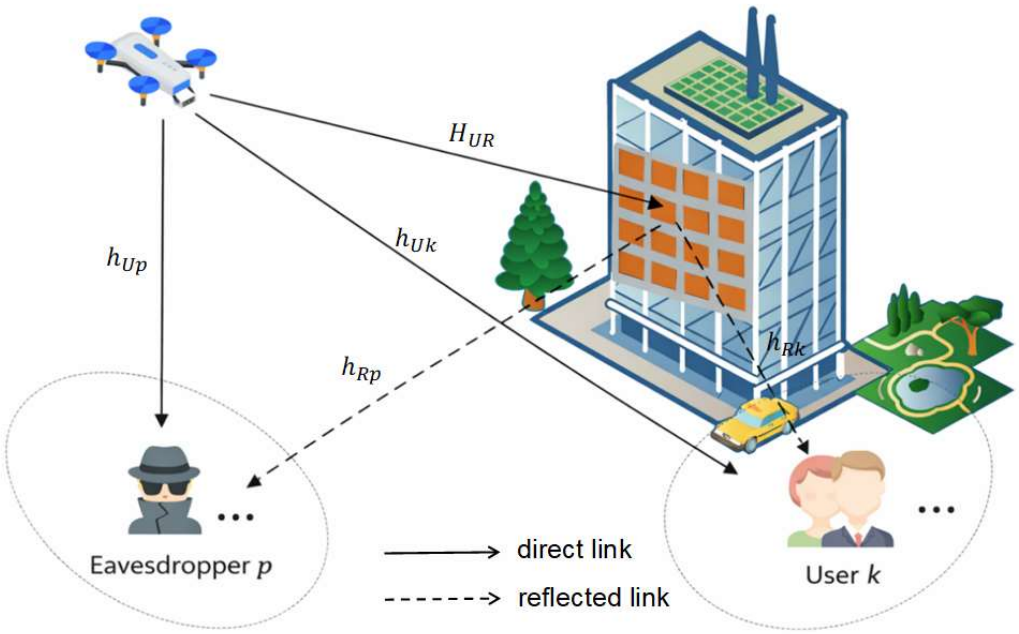

Fig. 1 RIS-Assisted UAV Communication System Model

2.2 UAV Jitter Model

Practical UAV flight is often affected by airflow disturbance and airframe vibration, which may cause random attitude jitter. In this paper, the UAV attitude jitter is modeled as a small-angle random rotational perturbation[9]. The jitter at time slot n is defined as

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

$$\delta[n] = [\delta_r[n], \delta_p[n], \delta_y[n]]^{\mathrm{T}} \sim \mathcal{N}(0, \sigma_{\mathrm{jit}}^2 \mathbf{I}_3) \tag{1}$$

where $\delta_r[n]$, $\delta_p[n]$, and $\delta_y[n]$ denote roll, pitch, and yaw perturbations, respectively.

Let $C^{\mathrm{nom}}[n]$ denote the UAV attitude matrix without jitter. Then, the actual attitude matrix under jitter is given by:

$$C^{\mathrm{true}}[n] = C^{\mathrm{nom}}[n] R(\delta[n]) \tag{2}$$

where $R(\delta[n])$ represents the rotation perturbation caused by attitude jitter.

In this work, the UAV position is assumed unchanged, while the antenna pointing direction is affected. This results in angle deviation and leads to array mismatch and CSI uncertainty.For the UAV–RIS link, the array response at the UAV side is expressed as

$$a_U(\mu_{UR}[n]) \triangleq \frac{1}{\sqrt{A}} \left[1, e^{j\pi\mu_{UR}[n]}, \ldots, e^{j\pi(A-1)\mu_{UR}[n]}\right]^T \in \mathbb{C}^{A\times 1} \tag{3}$$

Without jitter, the direction parameter is denoted as $\mu_{UR}^{nom}[n]$, while under jitter it becomes $\mu_{UR}^{true}[n]$. Therefore, attitude jitter leads to array-response deviation and imperfect CSI.

2.3 Transmission Model

In the considered mmWave scenario, the UAV–RIS link is assumed to be dominated by the line-of-sight path. Therefore, the channel matrix of the UAV–RIS link at time slot n is modeled as

$$H_{UR}[n] = \sqrt{\beta_{UR}(d_{UR}[n])} e^{-\frac{2\pi}{\lambda_c} d_{UR}[n]} a_R(\varphi_{UR}[n], \vartheta_{UR}[n]) a_U^H(\mu_{UR}[n]) \tag{4}$$

where λ_c is the carrier wavelength, $d_{UR}[n]$ is the distance between the UAV and the RIS, and $a_R(\cdot)$ and $a_U(\cdot)$ denote the array response vectors of the RIS and the UAV, respectively. The large-scale fading is modeled by a standard path-loss model. The RIS phase-shift matrix is written as

$$\mathbf{\Phi}[n] = \mathrm{diag}(e^{j\theta_1[n]}, \ldots, e^{j\theta_M[n]}) \tag{5}$$

where $\theta_m[n]$ denotes the phase shift of the m-th RIS element. For user or eavesdropper i, the equivalent channel is expressed as

$$h_{\mathrm{eff},i}^H[n] = h_{Ui}^H[n] + h_{Ri}^H[n] \mathbf{\Phi}[n] H_{\mathrm{UR}}[n] \tag{6}$$

Since the controller cannot directly observe the actual attitude jitter, the CSI used for optimization is constructed from the nominal UAV orientation, while the true channel is determined by the jittered orientation. Therefore, the nominal channel and the true channel are not identical, i.e.,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

$$\mathbf{H}_{\mathrm{UR}}^{true}[n] \neq \mathbf{H}_{\mathrm{UR}}^{\mathrm{nom}}[n] \tag{7}$$

This mismatch reflects the impact of UAV attitude jitter on practical channel estimation and system performance. At time slot n, the UAV adopts linear beamforming for multi-user transmission. The transmitted signal is given by

$$\mathbf{x}[n] = \mathbf{G}[n]s[n] = \sum_{k=1}^{K} g_k[n]s_k[n] \tag{8}$$

where $g_k[n]$ is the precoding vector for user k.Then, the received signal at receiver i is

$$y_i[n] = h_{\mathrm{eff},i}^H[n]\mathbf{x}[n] + n_i[n] \tag{9}$$

where $n_i[n]$ is additive Gaussian noise. Accordingly, the SINR of user k is given by

$$\gamma_k[n] = \frac{\left|h_{\mathrm{eff},k}^H[n]g_k[n]\right|^2}{\sum_{j\neq k}\left|h_{\mathrm{eff},k}^H[n]g_j[n]\right|^2 + \sigma_k^2} \tag{10}$$

and the achievable rate is

$$R_k[n] = \log_2(1+\gamma_k[n]) \tag{11}$$

Similarly, the eavesdropping rate for eavesdropper p to decode user k is denoted by $R_{p,k}[n]$. Therefore, the secrecy rate of user k is defined as

$$R_k^{\mathrm{sec}}[n] = \left[R_k[n] - \max_{p\in\mathcal{P}} R_{p,k}[n]\right]^+ \tag{12}$$

3. Problem Formulation

In this paper, the objective is to maximize the time-averaged sum secrecy rate over the considered time horizon by jointly optimizing the UAV trajectory Q, the active beamforming matrix G, and the RIS phase-shift matrix Φ. The optimization problem is formulated as

$$\begin{aligned}
(\mathrm{P1}) \quad & \max_{\mathbf{Q},\mathbf{G},\boldsymbol{\Phi}} \frac{1}{N}\sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} R_k^{\mathrm{sec}}[n] \\
\text{s.t.} \quad & \mathrm{C1} \quad \|\mathbf{q}[n+1]-\mathbf{q}[n]\|^2 \le D_{\max}, \quad \forall n = 0,\ldots,N-2 \\
& \mathrm{C2} \quad x_{\min} \le x_U[n] \le x_{\max}, \quad \forall n \in \mathcal{N} \\
& \qquad y_{\min} \le y_U[n] \le y_{\max}, \quad \forall n \in \mathcal{N} \\
& \mathrm{C3} \quad \mathrm{Tr}\left(\mathbf{G}[n]\mathbf{G}^H[n]\right) \le P_{\max}, \quad \forall n \in \mathcal{N} \\
& \mathrm{C4} \quad \boldsymbol{\Phi}[n] = \mathrm{diag}(\mathbf{v}[n]), \quad \forall n \in \mathcal{N} \\
& \qquad |v_m[n]| = 1, \quad \forall m = 1,\ldots,M,\, n \in \mathcal{N}
\end{aligned} \tag{13}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

The above constraints correspond to the UAV mobility limit, flight region, transmit power budget, RIS unit-modulus requirement, and the predefined initial UAV position, respectively. Since the UAV trajectory, beamforming, and RIS phase shifts are strongly coupled and the objective is non-convex, the problem is difficult to solve using conventional optimization methods. Therefore, a deep reinforcement learning based method is adopted in the following section.

4. Problem Solution

4.1 Dual-agent Framework

Inspired by[10], the considered optimization problem is addressed using a dual-agent deep reinforcement learning framework. Since the controller cannot directly observe the UAV attitude jitter and can only rely on the estimated CSI constructed from jitter-free pointing, the problem is inherently partially observable. For simplicity, it is approximated as a Markov Decision Process (MDP) during policy learning.

To reduce the complexity of the high-dimensional continuous optimization problem, two agents are designed to cooperate with each other. This design separates communication control from mobility control, thereby reducing the learning difficulty caused by the large continuous action space. Agent 1 is responsible for communication control, including the design of the beamforming matrix and RIS phase shifts. Agent 2 is responsible for mobility control, which updates the UAV trajectory in the horizontal plane. At each time slot, the two agents make decisions based on their respective observations and jointly interact with the environment to optimize system performance. For Agent 1, the observation consists of the estimated equivalent channels of all legitimate users and the eavesdropper. Based on this information, it generates continuous control variables for beamforming and RIS phase adjustment. For Agent 2, the observation mainly includes the UAV position, RIS location, user positions, and eavesdropper positions, based on which it determines the UAV movement at each time slot. In this way, the two agents focus on different control tasks while still pursuing a unified long-term optimization objective. To enhance secrecy performance while avoiding constraint violations, a shared reward mechanism is adopted. The reward at time slot n is defined as

$$r[n] = \tanh\left(\sum_{k \in \mathcal{K}} R_k^{\text{sec}}[n] - c_1 p_m[n] - c_2 p_r[n] - c_3 p_g[n] \right) \tag{14}$$

where $R_k^{\text{sec}}[n]$ denotes the secrecy rate of user k, $p_m[n]$ represents the maneuver constraint penalty, $p_r[n]$ denotes the flight-region penalty, and $p_g[n]$ represents the transmit power penalty. The coefficients c_1, c_2, and c_3 are weighting parameters. Through this reward design, Agent 1 learns robust communication strategies, while Agent 2 learns an effective UAV trajectory. Moreover, the shared reward encourages coordinated behavior between the two agents, enabling joint optimization of communication and mobility under the impact of UAV attitude jitter.

4.2 DRL Algorithm

To improve training stability and sample efficiency, several stabilization techniques are incorporated into the learning process of each agent, including twin critic networks, target policy smoothing, delayed policy updates, and prioritized experience replay (PER)[11]. For each agent, one actor network and two critic networks are constructed, together with their corresponding target networks. These techniques are introduced to mitigate instability issues in continuous control tasks and to improve the robustness of policy learning.

During training, mini-batch samples are drawn from the replay buffer according to the PER strategy. Target policy smoothing is introduced by adding clipped Gaussian noise to the target action, which helps reduce the sensitivity of policy learning to approximation errors. In addition, clipped double-Q

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

learning is employed to alleviate Q-value overestimation and improve training robustness. The corresponding target value is computed as

$$y = r + \gamma(1-d)\min(Q_1, Q_2) \tag{15}$$

where γ denotes the discount factor and d is the terminal indicator.

To further enhance sample utilization, PER is adopted so that more informative samples are selected more frequently during training. Moreover, delayed policy updates are applied, where the actor network is updated less frequently than the critic networks, which improves convergence stability. Finally, soft target updates are performed to gradually synchronize the target networks with the online networks, i.e.,

$$\theta_i^- \leftarrow \tau\theta_i + (1-\tau)\theta_i^- \tag{16}$$

where τ is the soft-update coefficient.

By integrating these mechanisms, the proposed dual-agent QA-DDPG framework achieves more stable and robust policy learning for the joint optimization of UAV trajectory, beamforming, and RIS phase shifts under CSI mismatch caused by UAV attitude jitter. Therefore, the proposed approach is well suited for handling the coupled control problem and channel uncertainty in RIS-assisted UAV secure communication systems.

5. Simulation Results

5.1 Parameter Settings

In this section, numerical simulations are conducted to evaluate the performance of the proposed RIS-assisted mmWave UAV secure transmission scheme under UAV attitude jitter.

A dual-agent actor–critic reinforcement learning framework is adopted for training, where each episode consists of 100 time slots and the total training process includes 100 episodes.

The main simulation parameters are set as follows: the number of RIS elements is $M = 16$, the number of UAV antennas is $A = 4$, the number of legitimate users is $K = 2$, and the number of eavesdroppers is $P = 1$. The maximum transmit power is $P_{\max} = 30$ dB, and the noise power is $\sigma^2 = -144$ dBm. The path-loss exponents are set to $\alpha_{UR} = 2.2$, $\alpha_{Ui} = 3.5$, and $\alpha_{Ri} = 2.8$. Unless otherwise specified, these parameters are used throughout the simulations.

5.2 Simulation Analysis

Fig. 2 illustrates the impact of UAV attitude jitter intensity on angular deviation and CSI mismatch. As the jitter standard deviation increases, both the angular deviation and the normalized CSI mismatch show a clear increasing trend. This is because UAV attitude jitter introduces small-angle rotational perturbations, which lead to deviations in the AoA/AoD and further cause array-response mismatch. As a result, a discrepancy between the estimated CSI and the true CSI is observed. These results confirm that UAV attitude jitter significantly affects channel accuracy and should be properly considered in system design. Therefore, robust optimization methods are necessary for RIS–UAV secure transmission under jitter conditions.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

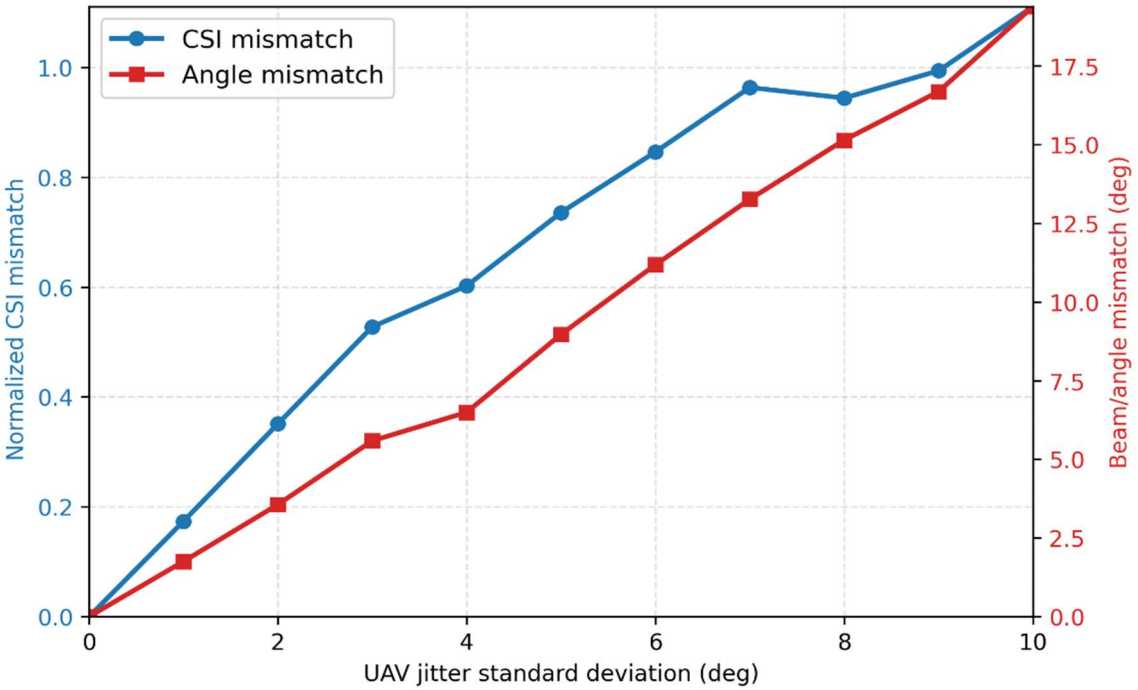

Fig. 2 Impact of jitter intensity on angular and CSI errors

Fig. 3 compares the average rate and secrecy rate of the proposed QA-DDPG algorithm with several baseline DRL methods, including TwinDDPG, TD3, and DDPG, under different UAV attitude jitter intensities. It can be observed that both the average rate and secrecy rate decrease as the jitter standard deviation increases, indicating that CSI mismatch caused by UAV attitude jitter degrades system performance. Notably, the proposed QA-DDPG algorithm consistently achieves the best performance across all jitter levels. In particular, under moderate and strong jitter conditions, it shows smaller performance degradation compared with the baseline methods, demonstrating its robustness against jitter-induced CSI uncertainty. In contrast, the conventional DDPG algorithm suffers from the most severe performance degradation, mainly due to its Q-value overestimation issue, which makes it more sensitive to time-varying CSI mismatch.

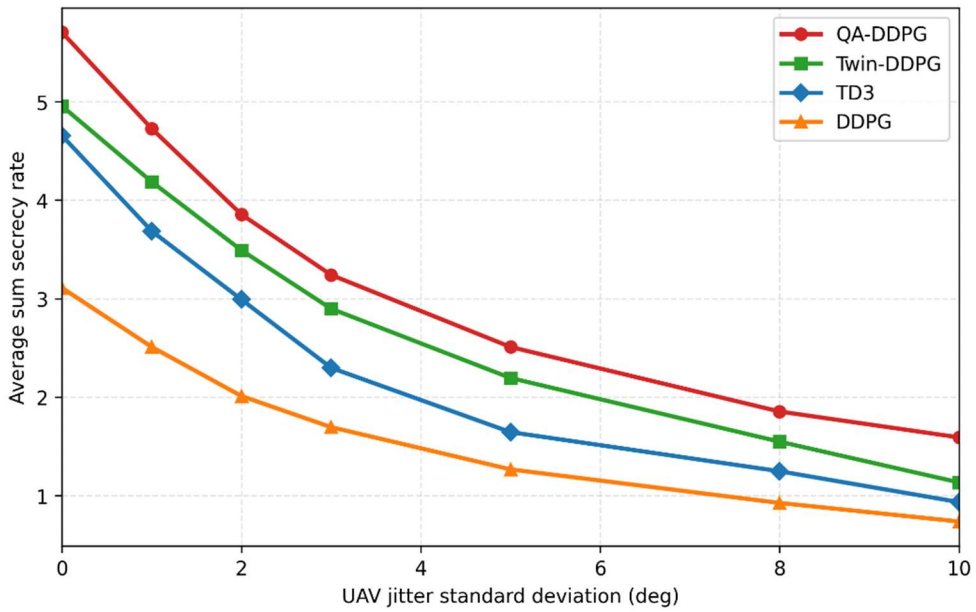

Fig. 3 Secrecy rate of different algorithms under varying jitter

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

Fig. 4 analyzes the performance of the proposed joint optimization scheme compared with several baseline methods with fixed variables, including fixed RIS phase shifts, fixed UAV trajectory, and fixed active beamforming. The results show that the proposed joint optimization consistently achieves the highest average rate and secrecy rate during the training process. This indicates that the UAV trajectory, active beamforming, and RIS phase shifts are strongly coupled, and joint optimization is necessary to fully exploit system performance. Among the baseline schemes, fixing the RIS phase shifts leads to the most significant performance degradation, highlighting the critical role of RIS in shaping the cascaded channels. In addition, fixing either the UAV trajectory or the active beamforming also results in noticeable performance loss, further confirming that optimizing a single component is insufficient for achieving optimal secure transmission.

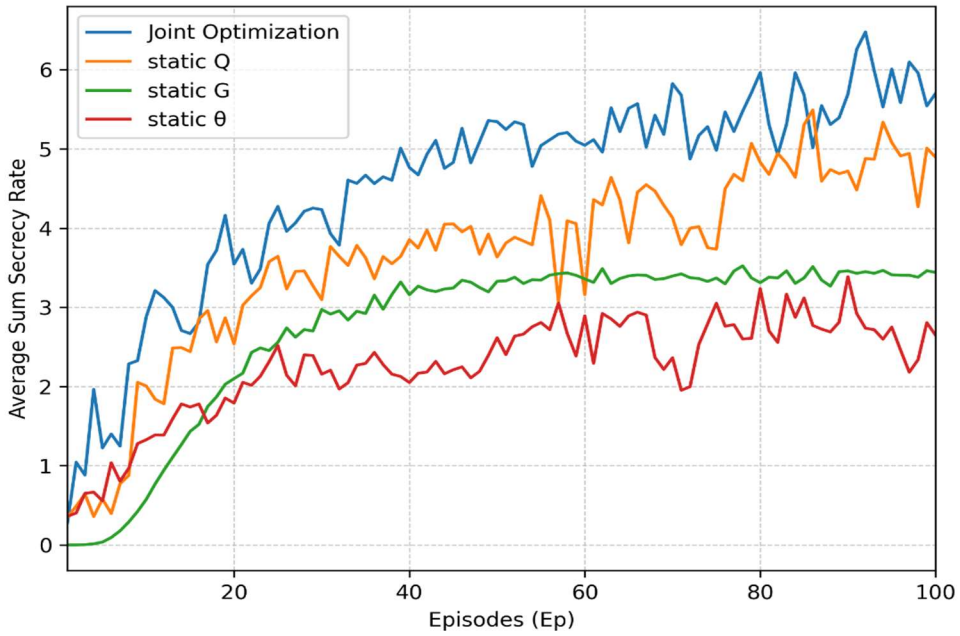

Fig. 4 Joint optimization vs. fixed baselines

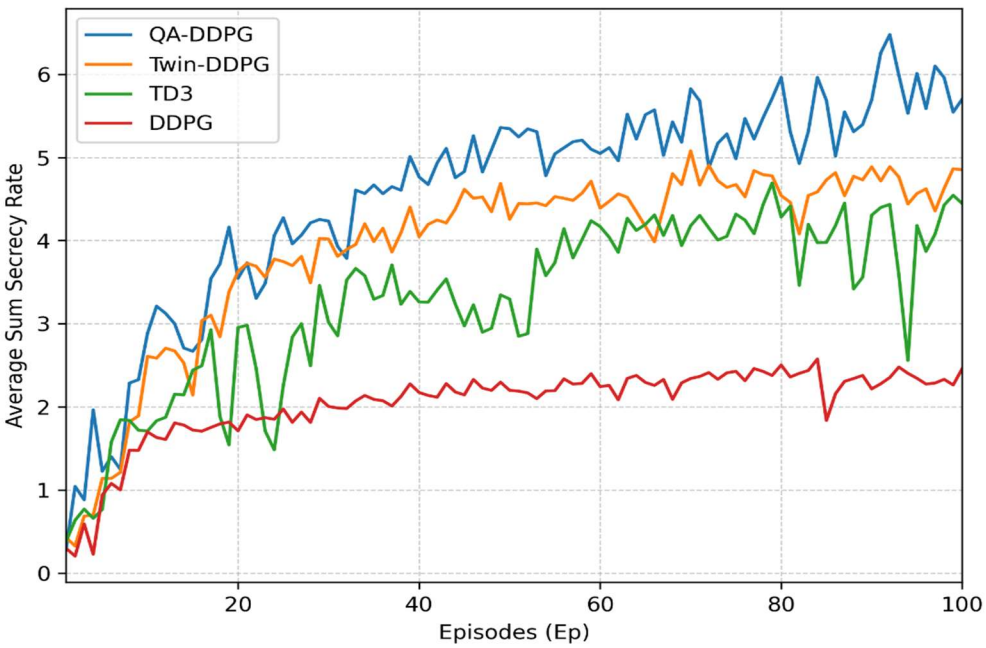

Fig. 5 Training convergence of different algorithms

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0015

Fig. 5 compares the training convergence behaviors of the QA-DDPG, TwinDDPG, TD3, and DDPG algorithms under UAV attitude jitter conditions. It can be observed that the proposed QA-DDPG algorithm achieves higher average and secrecy rates in the early training stage and gradually converges to a superior performance level with stable behavior. In contrast, the conventional DDPG algorithm shows the slowest convergence and the lowest final performance. Although TD3 improves the convergence speed compared with DDPG, it still exhibits noticeable oscillations during training.

These results demonstrate that the proposed QA-DDPG framework provides more stable and efficient learning under time-varying CSI mismatch caused by UAV attitude jitter.

6. Conclusion

This paper investigates robust joint optimization in an RIS-assisted mmWave UAV secure downlink system under UAV attitude jitter. The jitter is modeled to capture array-response mismatch and CSI uncertainty. To address the strongly coupled and non-convex optimization of UAV trajectory, active beamforming, and RIS phase shifts, a dual-agent reinforcement learning framework, termed QA-DDPG, is proposed. Simulation results show that UAV attitude jitter significantly degrades the secrecy performance due to angular deviation and CSI mismatch. The proposed QA-DDPG achieves higher secrecy rates and demonstrates strong robustness under different jitter intensities. In addition, it outperforms baseline methods in terms of convergence speed and training stability.

These results confirm the effectiveness of the proposed method and highlight the importance of joint optimization for RIS–UAV secure transmission under jitter conditions.

References

[1] Geraci G, Garcia-Rodriguez A, Azari M M, et al. What will the future of UAV cellular communications be? A flight from 5G to 6G[J]. IEEEcommunications surveys tutorials, 2022, 24(3): 1304-1335.

[2] Sun X, Ng D W K, Ding Z, et al. Physical layer security in UAV systems: Challenges and opportunities[J]. IEEE Wireless Communications, 2019, 26(5): 40-47.

[3] Jiang T, Yu W. Interference nulling using reconfigurable intelligent surface[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(5): 1392-1406.

[4] Khan W U, Lagunas E, Ali Z, et al. Opportunities for physical layer security in UAV communication enhanced with intelligent reflective surfaces[J]. IEEE Wireless Communications, 2022, 29(6): 22-28.

[5] Zhang S, Hao W, Sun G, et al. Joint beamforming optimization for active STAR-RIS-assisted ISAC systems[J]. IEEE Transactions on Wireless Communications, 2024, 23(11): 15888-15902.

[6] Wen Y, Chen G, Fang S, et al. RIS-assisted UAV secure communications with artificial noise-aware trajectory design against multiple colluding curious users[J]. IEEE Transactions on Information Forensics and Security,2024, 19: 3064-3076.

[7] Sarhan A Y, Abdullah O A, Al-Hraishawi H, et al. Reinforcement learning-driven secrecy energy efficiency maximization in RIS-enabled communication systems[J]. IEEE Access, 2025.

[8] Bansal A, Agrawal N, Singh K, et al. RIS selection scheme for UAVbased multi-RIS-aided multiuser downlink network with imperfect andoutdated CSI[J]. IEEE Transactions on Communications, 2023, 71(8):4650-4664.

[9] Saeed A K, Salim M M, Nasir A A, et al. Throughput Optimization in UAV-Mounted RIS under Jittering and Imperfect CSI via DRL[J]. arXiv preprint arXiv:2512.24773, 2025.

[10] Guo X, Chen Y, Wang Y. Learning-based robust and secure transmission for reconfigurable intelligent surface aided millimeter wave UAV communications[J]. IEEE Wireless Communications Letters, 2021, 10(8): 1795-1799.

[11] Brittain M, Bertram J, Yang X, et al. Prioritized sequence experiencereplay[J]. arXiv preprint arXiv:1905.12726, 2019.