

Physics-Based Trajectory Design for Cellular-Connected UAV in Rainy Environments Based on Deep Reinforcement Learning

Hao Qin¹, Zhaozhou Wu¹, and Xingqi Zhang^{1,2}

Abstract—Cellular-connected unmanned aerial vehicles (UAVs) have gained increasing attention due to their potential to enhance conventional UAV capabilities by leveraging existing cellular infrastructure for reliable communications between UAVs and base stations. They have been used for various applications, including weather forecasting and search and rescue operations. However, under extreme weather conditions such as rainfall, it is challenging for the trajectory design of cellular UAVs, due to weak coverage regions in the sky, limitations of UAV flying time, and signal attenuation caused by raindrops. To this end, this paper proposes a physics-based trajectory design approach for cellular-connected UAVs in rainy environments. A physics-based electromagnetic simulator is utilized to take into account detailed environment information and the impact of rain on radio wave propagation. The trajectory optimization problem is formulated to jointly consider UAV flying time and signal-to-interference ratio, and is solved through a Markov decision process using deep reinforcement learning algorithms based on multi-step learning and double Q-learning. Optimal UAV trajectories are compared in examples with homogeneous atmosphere medium and rain medium. Additionally, a thorough study of varying weather conditions on trajectory design is provided, and the impact of weight coefficients in the problem formulation is discussed. The proposed approach has demonstrated great potential for UAV trajectory design under rainy weather conditions.

Index Terms—Cellular-connected UAV, deep reinforcement learning, parabolic wave equation, rain attenuation, trajectory design.

I. INTRODUCTION

UNMANNED aerial vehicles (UAVs) are gaining increasing popularity in a variety of applications, such as surveillance, monitoring, and inspection. However, the limited range of traditional UAVs can be a significant obstacle to their widespread adoption [1]–[6]. To overcome this challenge, cellular-connected UAVs have emerged as an innovative wireless technology that integrates UAVs into cellular networks [7]–[11]. On one hand, dedicated UAVs can serve as communication relays or even aerial base stations (BSs) to provide wireless communications among BSs and users that are located at long distances or cannot be directly connected [12].

On the other hand, cellular networks can provide a reliable and secure communication link between UAVs and ground control stations. This allows operators to fly the UAVs over longer distances and in more challenging environments, without the need for visual contact. In particular, the UAVs are supported by the existing cellular BSs on the ground.

In complex meteorological conditions such as rain, cellular-connected UAVs can be utilized for various applications, such as weather monitoring, disaster response, and agricultural inspections. However, several critical challenges need to be addressed when integrating UAVs into existing cellular networks in rainy environments. Firstly, weak coverage regions exist in the sky because such BSs and conventional cellular networks are primarily designed to serve terrestrial user equipment (UE), and BS antennas are typically downtilted towards the ground. Therefore, the trajectories of the cellular-connected UAVs need to be well-designed to meet the communication requirements and mission specifications. Secondly, as aerial UEs, UAVs require efficient and site-specific trajectories, which necessitates the development of rapid and physics-based models that can extract detailed information about the surroundings and design an appropriate trajectory. However, statistical models require time-consuming and labor-intensive measurement campaigns in rainy environments, while deterministic models can be employed as an alternative. Thirdly, under rainy weather conditions, wireless communication systems can experience significant degradation in performance, which can have a direct impact on the optimal trajectory of UAVs. This highlights the necessity for accurate wave propagation models in rainy environments. The current techniques on cellular-connected UAV path planning do not account for rainy weather conditions. Addressing this limitation is crucial to ensure the reliability and effectiveness of UAV communication in various weather conditions and to enhance the overall performance of cellular-connected UAV systems. All of these factors motivate the development of physics-based trajectory design approach for cellular-connected UAVs in rainy environments, which is missing in the current literature.

A. Related Prior Work

Trajectory design for cellular-connected UAVs has been extensively investigated in recent years [13]–[17]. In [18], convex optimization and linear programming were utilized to find the optimal set of intermediate points and speed for a UAV. The UAV trajectory is optimized to minimize the

Manuscript received XX XX, XXXX; revised XX XX, XXXX; accepted XX XX, XXXX.

¹The authors are with the School of Electrical and Electronic Engineering, University College Dublin, Ireland (e-mail: hao.qin@ucdconnect.ie, zhaozhou.wu@ucdconnect.ie, xingqi.zhang@ucd.ie).

²The author is with the Department of Electrical and Computer Engineering, University of Alberta, Canada (e-mail: xingqi.zhang@ualberta.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier: XXXX

connection time constraint with the ground BSs. Graph theory and convex optimization have also been used to ensure that the UAV remains connected to at least one BS while minimizing travel time [19]. However, the above-mentioned conventional optimization-based UAV trajectory design methods often face practical limitations, as many of the optimization problems are non-convex and difficult to solve effectively.

Alternatively, machine learning (ML) techniques have emerged as a promising alternative for UAV trajectory design [20]–[23]. For instance, the Q-learning method has been used to design trajectories that maximize continuous connection time [24], while a dueling double Q network has been employed to optimize UAV trajectory [25]. Additionally, deep reinforcement learning (DRL) has also shown great promise in addressing complex decision-making problems associated with UAV trajectory design. For example, a DRL algorithm, based on echo state network cells, was developed for UAV path planning [26], and DRL was also utilized to enhance the ability of UAVs to understand the environment based on the captured image [27]. However, one of the main challenges associated with these ML techniques is the need for extensive measurement campaigns, which can be prohibitively expensive for large-scale scenarios or complex environments. Deterministic models such as ray tracing have been utilized to generate radio maps for UAV trajectory design, but these methods do not take into account factors such as rain attenuation, which can significantly affect the performance of cellular-connected UAVs.

Rain attenuation is a well-studied phenomenon in radio wave propagation, and several empirical and semi-empirical models have been proposed to predict the attenuation caused by rain [28], [29]. However, it should be noted that these models are primarily focused on point-to-point or point-to-multipoint attenuation, and do not account for the effects of multipath propagation. Therefore, there is a need for further research to explore and develop techniques that can accurately model the impact of multipath propagation on rain attenuation prediction. The parabolic wave equation (PWE) method has shown promise in modeling wave propagation, as it can properly take into account wave refraction and diffraction [30]–[32]. Additionally, the PWE method has proved to be effective in predicting radio wave propagation under complex meteorological conditions [33].

B. Contributions

Motivated by the above facts, we propose a deep reinforcement learning-based framework, which can design trajectories for cellular-connected UAVs in rainy environments using rainfall parameters and detailed information about the environment's geometry. The main contributions of this paper are summarized as follows:

- Firstly, we formulate the UAV trajectory optimization problem to minimize the difference between the mission completion time and the weighted signal-to-interference ratio (SIR). We then transform this problem into a sequential decision-making one, and reformulated it as a Markov decision process (MDP) with a designed state space, action space, and reward function.
- Secondly, we employ a physics-based deterministic wave propagation model to extract the SIR, replacing the time-consuming measurement process. This model allows us to consider the detailed information about the environment's geometry and the impact of meteorological conditions on radio wave propagation and trajectory design.
- Thirdly, we utilize deep reinforcement learning to address the reformulated MDP problem, and this technique can learn policies in MDP. Besides, deep neural networks (DNN) are used to approximate the Q-function in DRL. We employ a dueling network architecture with a double deep Q-network (DDQN) to train the DNN. This architecture uses separate sets of weights to select actions and evaluate their Q-values, which reduces overestimation and improves stability.
- Finally, we conduct extensive evaluations of our proposed approach and investigate the impact of various rainy weather conditions on UAV trajectory design. We compare the optimal trajectory for cellular-connected UAVs in the atmosphere medium and rain medium, highlighting the necessity of accurate physics-based wave propagation models for UAV trajectory design. We also present a comparative analysis of different rainfall parameters and weight coefficients in the problem. Our results demonstrate the effectiveness of our approach in managing varying weather conditions, and we show that weight coefficients can be adjusted based on the priorities of the users, whether they are more concerned with flying time or SIR.

The rest of this paper is organized as follows. In Section II, we present the system model and formulate the trajectory design problem. Section III provides an introduction to the wave propagation model in rainy environments, while Section IV gives an overview of deep Q learning. In Section V, we present our proposed DRL-based trajectory design, while the numerical results and performance evaluation of our approach are given in Section VI. Finally, we conclude the paper in Section VII.

II. SYSTEM MODEL AND PROBLEM FORMULATION

As shown in Fig. 1, a cellular-connected UAV system is considered, where UAV is supported by cellular BSs. The considered airspace can be represented as a cubic volume, which can be specified by $X \times Y \times Z$, where $X \in [x_L, x_H]$, $Y \in [y_L, y_H]$, and $Z \in [z_L, z_H]$.

We assume that the UAV takes off from a random initial location, \mathbf{q}_I , and flies towards a fixed final destination, \mathbf{q}_F . If a fixed height is assumed for the UAV's flight, we can visualize the range of movement for the UAV as Fig. 2, where (x_L, y_L) and (x_H, y_H) are the boundary points of the considered airspace at the fixed height. An optimal trajectory can be designed for the UAV to minimize its flying time while maintaining reliable communication connectivity with BSs. The UAV trajectory can be represented by $\mathbf{q}(t), t \in [0, T]$, where T denotes the flying time. Hence, we have:

$$\mathbf{q}(0) = \mathbf{q}_I, \quad (1)$$

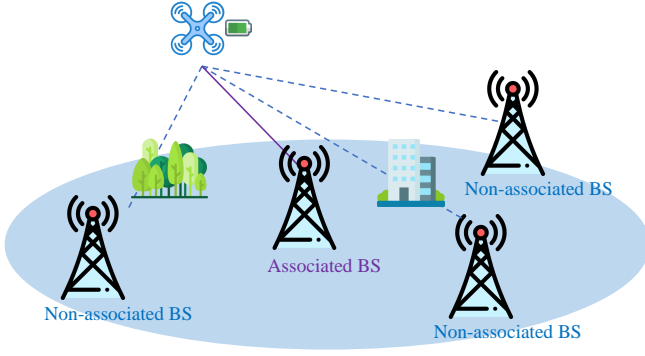


Fig. 1. Cellular-connected UAV communication systems.

$$\mathbf{q}(T) = \mathbf{q}_F, \quad (2)$$

$$\mathbf{q}_L \preceq \mathbf{q}(t) \preceq \mathbf{q}_H, \forall t \in (0, T). \quad (3)$$

where \preceq indicates the element-wise inequality. The coordinates of \mathbf{q}_L and \mathbf{q}_H are (x_L, y_L, z_L) and (x_H, y_H, z_H) , respectively.

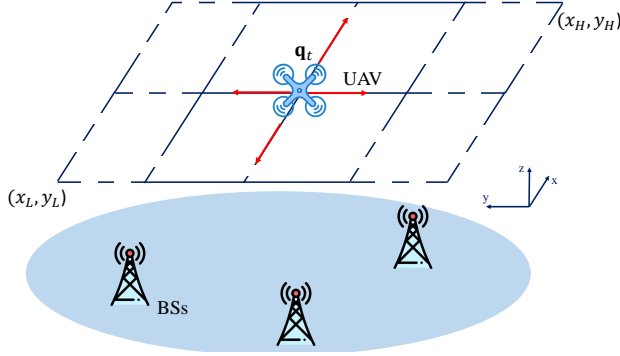


Fig. 2. Illustration of available grid points in the airspace.

In particular, the whole considered airspace is divided into M cells, and $u_m(t)$, $1 \leq m \leq M$, represents the received signal strength (RSS) from cell m to the UAV at time t . RSS in the rain medium can be generated by the PWE simulator, and a detailed description of the PWE method is provided in Section III. As the UAV is associated with one BS at a time, we define $b(t) \in \{1, \dots, M\}$ to denote the cell which is associated with the UAV at time t . Then, SIR at time t can be defined as:

$$\text{SIR}(\mathbf{q}(t), b(t)) = \frac{[u_{b(t)}(t)]^2}{\sum_{m \neq b(t)} [u_m(t)]^2}. \quad (4)$$

The total expected SIR along the UAV path can be expressed as:

$$\text{SIR}_{\text{total}}(\mathbf{q}(t), b(t)) = \int_0^T \text{SIR}(\mathbf{q}(t), b(t)) dt. \quad (5)$$

Intuitively, the UAVs can easily avoid weak coverage regions of the cellular network in the sky with longer flying times T . However, practical applications require reducing the

flying time due to the limited endurance of UAVs. Hence, there should be a tradeoff between minimizing the UAV flying time and maximizing the total SIR along the trajectory, which can be balanced by designing $\{\mathbf{q}(t)\}$ and $\{b(t)\}$. The trajectory design problem can be formulated as follows by introducing a certain weight μ :

$$(\text{P1}) : \max_{T, \{\mathbf{q}(t), b(t), \vec{\nabla}(t)\}} -T + \mu \text{SIR}_{\text{total}}(\mathbf{q}(t), b(t)) \quad (6)$$

$$\text{s.t.} \|\mathbf{q}(t+1) - \mathbf{q}(t)\| \leq V_{\max} \vec{\nabla}(t), \forall t \in [0, T-1], \quad (6a)$$

$$\mathbf{q}(0) = \mathbf{q}_I, \quad (6b)$$

$$\mathbf{q}(T) = \mathbf{q}_F, \quad (6c)$$

$$\mathbf{q}_L \preceq \mathbf{q}(t) \preceq \mathbf{q}_H, \forall t \in [0, T], \quad (6d)$$

$$b(t) \in \{1, \dots, M\}, \quad (6e)$$

$$\|\vec{\nabla}(t)\| = 1, \forall t \in [0, T], \quad (6f)$$

$$\text{SIR}_{\min} \leq \text{SIR}(\mathbf{q}(t), b(t)), \quad (6g)$$

where V_{\max} denotes the maximum UAV speed and $\vec{\nabla}(t)$ is the UAV flying direction. SIR_{\min} indicates the minimum SIR along the UAV path. Its specific definition can be tailored to suit the requirements of the actual applications being considered. Note that the weight μ is a non-negative coefficient. A higher value of the parameter μ indicates that the algorithm prioritizes cellular connectivity with the UAV over the flying time.

However, in practice, the optimization problem (P1) is highly non-convex, making it challenging to be efficiently solved. In addition, obtaining SIR in rain medium through measurement campaigns is a time-consuming and labor-intensive process. To address these issues, we propose a solution that employs a physics-based wave propagation simulator, PWE, along with a DRL algorithm based on a dueling deep network with multi-step learning.

III. PHYSICS-BASED WAVE PROPAGATION MODEL IN RAIN MEDIUM

The parabolic wave equation method [30] is a physics-based wave propagation modeling approach and can be expressed as

$$\frac{\partial u}{\partial z} = \frac{1}{2jk_0} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u. \quad (7)$$

In (7), u is the reduced plane wave solution, and radio waves are assumed to propagate along the z -axis. In general, (7) can be solved using different techniques, such as the finite difference method and the split-step Fourier method. It is found that the split-step Fourier scheme is more time-efficient compared to the finite difference scheme with the frequency increases, which makes it a good candidate for high-frequency propagation modeling [34]. Hence, in the paper, (7) is solved using a split-step Fourier technique as

$$u(x, y, z + \Delta z) = e^{-jk_0(n^2-1)\Delta x/2} \times \mathcal{F}^{-1} \{ C(k_x) C(k_y) \mathcal{F} \{ u(x, y, z) \} \}, \quad (8)$$

where \mathcal{F} and \mathcal{F}^{-1} are the Fourier transform pairs; k_x and k_y are the spectral variables, and

$$C(k) = \exp\left(\frac{-jk^2 \Delta z}{2k_0}\right). \quad (9)$$

In (8), $n = \sqrt{\varepsilon_{\text{eff}}}$ is the complex refractivity that can be calculated accordingly for the medium, and ε_{eff} is the relative effective permittivity of the medium.

The effective permittivity ε_{eff} can be defined as

$$\varepsilon_{\text{eff}} = \varepsilon_0 + \frac{\bar{P}}{\bar{E}}, \quad (10)$$

where ε_0 is the permittivity of the air, \bar{E} is the field in the medium, and \bar{P} denotes the average polarization density of the raindrops. In particular, \bar{P} can be calculated based on the raindrop spectrum $N(D)$, the polarization rate α , and the field \bar{E}_e inside the raindrops [35]:

$$\bar{P} = \int_{D_{\min}}^{D_{\max}} N(D) \alpha(D) \bar{E}_e(D) dD, \quad (11)$$

where D is the equivalent diameter of a single raindrop, and $\bar{E}_e(D)$ is the field inside raindrop. If the raindrop is a sphere:

$$\bar{E}_e = \bar{E} + \frac{\bar{P}}{3\varepsilon_0}. \quad (12)$$

If raindrop is an ellipsoid:

$$\bar{E}_e = \bar{E} + \frac{L_i \bar{P}}{\varepsilon_0}. \quad (13)$$

L_i denotes the polarization factor, which can be expressed as:

$$L_a = \frac{1}{e^2} \left(1 - \sqrt{\frac{1-e^2}{e^2}} \arcsin e \right), \quad (14)$$

$$L_b = L_c = \frac{1}{2} - \frac{1}{2} L_a, \quad (15)$$

where e is the eccentricity of an ellipsoidal raindrop that can be found in

$$e = \sqrt{1 - \left(\frac{a}{b} \right)^2}, \quad (16)$$

The rain medium is generally formed by raindrops of various shapes and sizes in the atmosphere. Since the drops with a diameter larger than 8 mm are unstable and can easily break up [32], we assume that the range for the diameter of raindrops is from 0.1 mm to 8 mm. If the diameter of the raindrop is more than 1.25 mm, the shape of the raindrop is a flat rotary ellipsoid. Otherwise, the shape can be assumed to be spherical. For the polarization rate α , if raindrop is a sphere,

$$\alpha(D) = 4\pi \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega - 2\varepsilon_0} \left(\frac{D}{2} \right)^3. \quad (17)$$

If raindrop is an ellipsoid,

$$\alpha(D) = v \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega + L_i(\varepsilon_\omega - \varepsilon_0)}, \quad (18)$$

where v is the volume of one raindrop, and ε_ω is the permittivity of water, which can be calculated by the Debye

formula [36]. Substituting (11) - (18) into (10), we can obtain the effective permittivity in rain medium:

$$\begin{aligned} \varepsilon_{\text{eff}} = & \varepsilon_0 + \int_{D_{\min}}^{1.25} N(D) 4\pi \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega - 2\varepsilon_0} \left(\frac{D}{2} \right)^3 \\ & \times \frac{3\varepsilon_0}{3\varepsilon_0 - 4\pi \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega - 2\varepsilon_0} \left(\frac{D}{2} \right)^3} dD \\ & + \frac{1}{3} \sum_{i=1}^3 \int_{1.25}^{D_{\max}} N(D) \frac{4\pi}{3} \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega + L_i(\varepsilon_\omega - \varepsilon_0)} \left(\frac{D}{2} \right)^3 \\ & \times \frac{\varepsilon_0}{\varepsilon_0 - L_i \frac{4\pi}{3} \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega + L_i(\varepsilon_\omega - \varepsilon_0)} \left(\frac{D}{2} \right)^3} dD. \end{aligned} \quad (19)$$

where the raindrop spectrum is chosen as the Marshall-Palmer spectrum, which is a function of the rain rate R (mm/h).

For validation purposes, as shown in Fig. 3, we compare the numerical results obtained from the PWE-based electromagnetic simulator at different frequencies against the ITU-R model [37], which is a widely used statistical model for predicting the effects of rain on radio wave propagation. The attenuation of the ITU-R model depends on the rain rate R :

$$\gamma_R = kR^\alpha, \quad (20)$$

where the coefficients k and α can be found in [37].

In this study, we set the rain rate as $R = 12.5$ mm/h and compared the attenuation results obtained from the PWE simulator and the ITU-R model at different frequencies, as shown in Fig. 3. The results demonstrate a good match between the PWE simulator and the ITU model, indicating the reliability of using the PWE simulator in rainy environments.

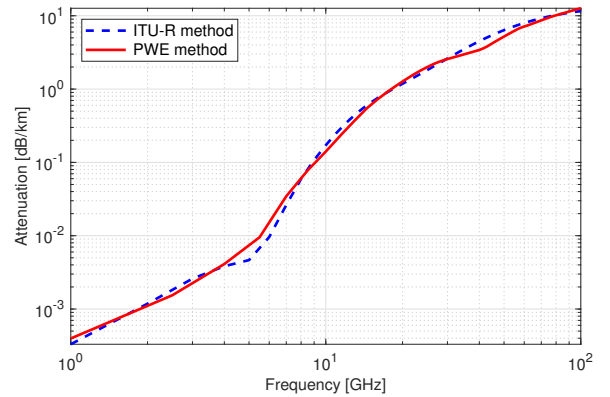


Fig. 3. Rain attenuation obtained from the PWE simulator and the ITU-R model at different frequencies.

It's worth noting that the ITU-R model relies on a large database of experimental measurements to provide empirical relationships, but it has limitations in complex environments where terrain and obstacles can significantly affect wave propagation. Additionally, the ITU-R model cannot provide a 3D received signal strength distribution in a given skyspace. Therefore, in our proposed approach, we employed the PWE simulator to obtain the electromagnetic wave propagation characteristics in rainy environments.

IV. PRELIMINARIES: DEEP Q NETWORK

In this section, we provide a brief overview of the deep Q network (DQN). A more comprehensive description can be found in [38]. DQN is a widely adopted reinforcement learning method designed to solve MDP in continuous state spaces. In DQN, an agent interacts with the environment to collect data to train itself. Unlike basic reinforcement learning in tabular settings, which stores the expected return of each action under a specific state in tabular form, DQN parameterizes the expected return of each action under a specific state using deep neural networks (DNNs). As a result, the policy of the DQN agent is governed by the parameters of DNN, which are updated using gradient descent with respect to the expected returns calculated using *bootstrapping* [38].

Specifically, at the start of the training, a DQN network with random coefficients θ is initialized. This network is referred to as the Q network. At each state $s \in \mathcal{S}$, the Q network selects the action that maximizes its output, known as the *exploitation* of the known expected returns predicted by the Q network. Additionally, with a probability of ϵ , the Q network selects a random action from the available actions, known as the *exploration*. This allows for unexplored action-state pairs to be updated in the current Q network. The parameter ϵ is used to balance the exploitation and exploration and decreases as a function of training episodes. This policy for selecting an action is known as the ϵ -greedy policy and is shown as follows:

$$f(a, b) = \begin{cases} a \leftarrow a_{id}, id \sim \mathcal{U}\{1, |\mathcal{A}|\} & \text{with probability } \epsilon, \\ a \leftarrow \underset{a}{\operatorname{argmax}} \hat{Q}(s, a; \theta), a_{id} \in \mathcal{A} & \text{with probability } 1 - \epsilon. \end{cases} \quad (21)$$

Once an action is selected according to (21), the next state s_{next} can be obtained using current state s , action a and transition model $P(s_{next}|s, a)$. In our settings, the transition is deterministic, specifically, $s_{next} = s + a$ with $P(s_{next}|s, a) = 1$. Before moving to the next state, the reward R at the current is calculated. The (s, R, a, s_{next}) tuple is the data that the DQN agent collected in the training phase for the training of the DNN. Such a tuple is known as a *transition*. With such transitions, expected returns of action a at state s can be estimated using bootstrapping and Monte Carlo (MC) methods. Specifically, transitions are stored in a replay memory D with capacity C . Then, a minibatch of B transitions is randomly sampled from D with $B \ll C$. The expected returns of the sample transitions are estimated using bootstrapping in (22), where γ is the discount factor. (22) uses the next action that maximizes the expected return predicted by the Q network state s_{next} instead of the *actual action* taken at s_{next} , which is a method used in Q learning to allow quick convergence.

$$\mathbb{E}(Q_{s,a}) = \begin{cases} R & \text{if episode ends,} \\ R + \gamma \max_{a \in \mathcal{A}} \hat{Q}(s_{next}, a; \theta) & \text{else.} \end{cases} \quad (22)$$

The DNN parameters θ can be trained using gradient descent with respect to the expected returns. The loss function is formulated in (23):

$$L(\theta) = \|\hat{Q}(s, a; \theta) - \mathbb{E}(Q_{s,a})\|. \quad (23)$$

However, $\mathbb{E}(Q_{s,a})$ in (22) is known to be a noisy estimation with high variance. This is because, in (22), the estimation itself depends on the parameters that need to be trained. A few techniques are used in addition to the above DQL to further stabilize and speed up the training process. These include the *target network* [39] and the *multi-step learning* [40]. Target network suggests the maintenance of a *target network* which is updated after multiple episodes. The target network shares the same architecture as the DNN. It is used for the prediction of the expected returns for different actions at s_{next} . With a target network, the estimation of the expected returns in (22) becomes:

$$\mathbb{E}(Q_{s,a}) = \begin{cases} R & \text{if episode ends,} \\ R + \gamma \max_{a \in \mathcal{A}} \hat{Q}(s_{next}, a; \theta^-) & \text{else.} \end{cases} \quad (24)$$

In (24), $\mathbb{E}(Q_{s,a})$ becomes independent of the DNN parameters θ and the training process becomes more stable. The parameters θ^- of the target network are set to be updated after a few training episodes with $\theta^- \leftarrow \theta$.

Multi-step learning provides more efficient use of the replay memory and a more stable training process by learning from multi-step bootstrap targets using a sliding window W (a queue) with length N_1 . N_1 transitions are stored in W before being stored in the replay memory. The transitions sliding window are used to calculate the *multi-step reward*. Specifically, for state-action pair (s, a) , its multi-step reward $R_{(0:N_1)}$ is calculated using (25) with $n = 0$ ((s, a) as the current state).

$$R_{n:(n+N_1)} = \sum_{i=0}^{N_1-1} \gamma^i R_{n+i+1}. \quad (25)$$

Then $(s_n, a_n, R_{(0:N_1)}, s_{n+N_1})$ is stored as a transition into the replay memory D . After the calculation of $R_{(0:N_1)}$, (s, a) is popped out of the head of W , and a new transition is appended to the end of W . Now the calculation of $\mathbb{E}(Q_{s,a})$ becomes:

$$\mathbb{E}(Q_{s,a}) = \begin{cases} R_0 & \text{if episode ends,} \\ R_{(0:N_1)} + \gamma \max_{a \in \mathcal{A}} \hat{Q}(s_{next}, a; \theta^-) & \text{else.} \end{cases} \quad (26)$$

In this paper, a dueling DQN [41] is adopted for the DNN architecture. Compared to the standard DQN, dueling DQN is able to learn the state-value function more efficiently.

V. PROPOSED TRAJECTORY DESIGN APPROACH FOR CELLULAR-CONNECTED UAVS

In this section, we present our proposed DRL-based approach for solving the trajectory design problem of cellular-connected UAVs in rainy environments.

A. Reformulation as an MDP

To solve the trajectory design problem (P1) using reinforcement learning methods, the problem can be first reformulated as an MDP, where the future states are only affected by the current state of the system and the selected action of the agent. First, the flying time T can be separated into N discrete time steps. We use Δt to present the time interval and $T = N\Delta t$. Subsequently, the UAV trajectory $\mathbf{q}(t)$ can be represented by a sequence $\{\mathbf{q}_1, \dots, \mathbf{q}_N\}$. Hence, (6a) - (6d) can be rewritten as:

$$\|\mathbf{q}_{n+1} - \mathbf{q}_n\| \leq V_{\max} \vec{\mathbf{v}}_n, \forall n, \forall t \in [0, T-1], \quad (27)$$

$$\mathbf{q}_0 = \mathbf{q}_I, \quad (28)$$

$$\mathbf{q}_N = \mathbf{q}_F, \quad (29)$$

$$\mathbf{q}_L \preceq \mathbf{q}_n \preceq \mathbf{q}_H, \forall n. \quad (30)$$

Similarly, the association policy $b(t)$ can be represented as b_n . Besides, the UAV flying direction is approximated as $\vec{\mathbf{v}}_n = \vec{\mathbf{v}}(n * \Delta t)$ at time t . As a result, (6e) and (6f) can be written as:

$$b_n \in \{1, \dots, M\}, \quad (31)$$

$$\|\vec{\mathbf{v}}_n\| = 1, \forall n. \quad (32)$$

For each step n , SIR in rain medium can be calculated by the PWE simulator introduced in Section III. The location \mathbf{q}_n , the association policy b_n , and the information of the environment are all taken into account. Based on the above discussions, the problem (P1) can be approximated as follows:

$$(P2): \max_{N, \{\mathbf{q}_n, b_n, \vec{\mathbf{v}}_n\}} -N + \mu \sum_{n=1}^N \text{SIR}(\mathbf{q}_n, b_n) \quad (33)$$

$$\text{s.t.} \|\mathbf{q}_{n+1} - \mathbf{q}_n\| \leq V_{\max} \vec{\mathbf{v}}_n, \forall n, \quad (33a)$$

$$\mathbf{q}_0 = \mathbf{q}_I, \quad (33b)$$

$$\mathbf{q}_N = \mathbf{q}_F, \quad (33c)$$

$$\mathbf{q}_L \preceq \mathbf{q}_n \preceq \mathbf{q}_H, \forall n, \quad (33d)$$

$$b_n \in \{1, \dots, M\}, \quad (33e)$$

$$\|\vec{\mathbf{v}}_n\| = 1, \forall n, \quad (33f)$$

$$\text{SIR}_{\min} \leq \text{SIR}_n, \forall n. \quad (33g)$$

Furthermore, the MDP can be formulated as follows:

1) *States*: The state space is constructed of the potential UAV location, \mathbf{q}_n , in the considered airspace. It is defined as

$$\mathcal{S} = \{\mathbf{q} : \mathbf{q}_L \preceq \mathbf{q} \preceq \mathbf{q}_U\}. \quad (34)$$

2) *Actions*: The action space corresponds to the UAV's flying direction. For each time step, the UAV carries out an action. The action space can be defined as

$$\mathcal{A} = \{\vec{\mathbf{v}} : \|\vec{\mathbf{v}}_n\| = 1\}. \quad (35)$$

3) *Transition Probabilities*: The transition probabilities in this problem are deterministic and governed by

$$\|\mathbf{q}_{t+1} - \mathbf{q}_t\| \leq V_{\max} \vec{\mathbf{v}}_n, \forall n. \quad (36)$$

4) *Rewards*: The reward function consists of two parts, the flying time and the accumulated SIR along the path. The flying time is defined as the time taken for the UAV to move from its current position to the next position after taking action a . The SIR component of the reward is defined to encourage the UAV to avoid weak coverage regions. Specifically, if the UAV enters a location with a small SIR, a penalty with weight μ is incurred. The SIR penalty component of the reward is given by $\mu \text{SIR}(\mathbf{q}, b)$. The reward is defined as follows:

$$R(\mathbf{q}, b) = -1 + \mu \text{SIR}(\mathbf{q}, b). \quad (37)$$

After being formulated as the above MDP, the problem can be solved by applying a dueling DDQN with multi-step learning. By utilizing the dueling DDQN architecture, we mitigate the risk of suboptimal policies caused by overestimation bias in Q-value estimates, resulting in more stable learning. Moreover, the integration of double Q-learning further refines value estimation, leading to a more robust action selection. The adoption of multi-step learning boosts efficiency and sample utilization, accelerating the learning process and enabling the UAV to discover better navigation strategies.

B. Dueling DDQN Learning for Trajectory Design

In our proposed approach, we focus on the scenario in which the UAV flies in a horizontal plane with a fixed altitude. We use DNN to approximate the action-value function $Q(\mathbf{q}, \vec{\mathbf{v}})$. Such a function contains the state input and the action input. We discretize the action space \mathbf{A} into $K = 4$ values, and $\mathbf{A} = \{\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(K)}\}$. The state space \mathbf{S} is continuous. In this paper, the action-value function $Q(\mathbf{q}, \vec{\mathbf{v}})$ is approximated using the dueling DQN, as illustrated in Fig. 4, where θ denotes the coefficients of the dueling DQN to ensure that the output $\hat{Q}(\mathbf{q}, \vec{\mathbf{v}}; \theta)$ gives a good approximation to the true action-value function $Q(\mathbf{q}, \vec{\mathbf{v}})$. At each time step n , the input of such network is the location of the UAV, \mathbf{q}_n . K outputs correspond to different actions in the action space, \mathbf{A} . The multi-step DQN learning is to minimize the loss in (38):

$$(R_{n:(n+N_1)} + \gamma^{N_1} \max_{\mathbf{v}' \in \mathbf{A}} \hat{Q}(\mathbf{q}_{n+N_1}, \vec{\mathbf{v}}'; \theta^-) - \hat{Q}(\mathbf{q}_n, \vec{\mathbf{v}}_n; \theta))^2 \quad (38)$$

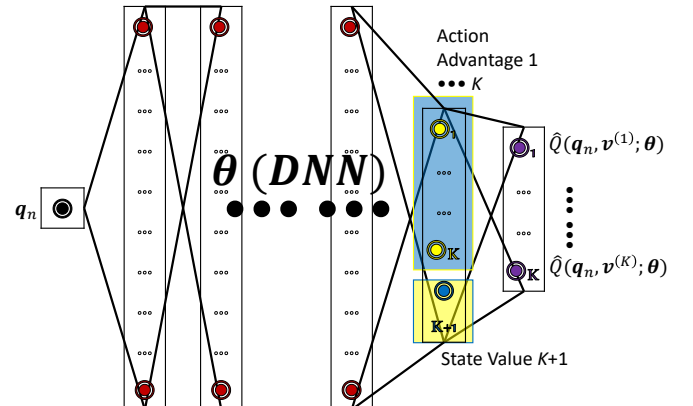


Fig. 4. Diagram of the dueling DQN for the trajectory design problem.

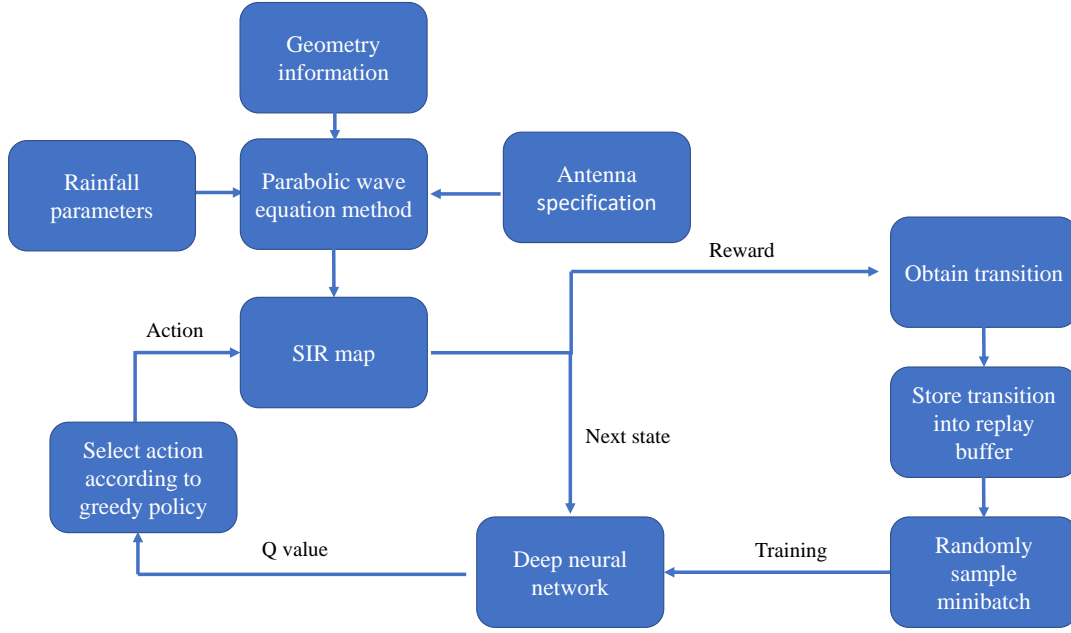


Fig. 5. Framework of the proposed approach of trajectory design for cellular-connected UAVs in rainy environments.

For multi-step DDQN learning, a truncated N_1 -step return in (39) from a given state \mathbf{q}_n is defined.

$$R_{n:(n+N_1)} = \sum_{i=0}^{N_1-1} \gamma^i R_{n+i+1}. \quad (39)$$

Furthermore, the loss of the multi-step learning in DDQN is

$$(R_{n:(n+N_1)} + \gamma^{N_1} \hat{Q}(\mathbf{q}_{n+N_1}, \vec{\mathbf{v}}^*; \boldsymbol{\theta}^-) - \hat{Q}(\mathbf{q}_n, \vec{\mathbf{v}}_n; \boldsymbol{\theta}))^2, \quad (40)$$

where

$$\vec{\mathbf{v}}^* = \underset{\vec{\mathbf{v}}' \in \mathcal{A}}{\text{argmax}} \hat{Q}(\mathbf{q}_{j+N_1}, \vec{\mathbf{v}}'; \boldsymbol{\theta}). \quad (41)$$

Coefficients $\boldsymbol{\theta}^-$ are used to evaluate the bootstrapping action in (40).

The proposed approach for trajectory design in rain medium is summarized in Algorithm 1, and its framework is illustrated in Fig. 5. The SIR map is calculated by the PWE simulator in Section III, and dueling DDQN multi-step learning is used. A sliding window queue with capacity N_1 is used to store the N_1 latest transitions. In particular, setting the maximum number of episodes \bar{N}_{epi} is an important step. This value determines the maximum number of iterations that the agent will interact with the environment and collect data to train the neural network. Once this maximum number of episodes is reached, the training process will stop, and the final DNN parameters will be saved as the trained model. Typically, the maximum number of episodes is chosen based on the complexity of the problem and the available computational resources. A larger number of episodes will generally lead to better performance, but it will also require more time and computational resources. On the other hand, a smaller number of episodes may not be sufficient to learn the optimal policy. Therefore, we also investigate the numerical results with various episodes in Section VI.

C. Computational Complexity Analysis

The problem addressed by dueling DDQN is defined by a continuous state space and a discrete action space with 4 possible actions. During the training process, the agent interacts with the environment, receiving rewards based on its actions and transition probabilities. The computational complexity of the approach primarily stems from two key components: the DNN utilized to approximate the action-value function and the multi-step learning process.

The computational complexity for computing the action-value function for a single state-action pair through DNN is approximately $O(T * (d + k))$, where d represents the depth (number of layers) of the network, k is the number of steps used for the multi-step update, and T is the number of time steps per episode. Additionally, the multi-step learning technique introduces a computational complexity of $O(b * d)$, where b is the size of each mini-batch. Overall, we estimate the total training computational complexity as $O(N * T * (d + k) + N * b * d)$, where N is the number of episodes.

VI. NUMERICAL RESULTS AND PERFORMANCE EVALUATION

To evaluate the effectiveness of our proposed approach, in this section, we present numerical results from a $2 \text{ km} \times 2 \text{ km}$ airspace. During testing, the UAV maintains a fixed altitude of 100 m while flying. The considered scenario includes seven ground-based base stations, whose locations are illustrated in Fig. 6. We generate the signal-to-interference ratio (SIR) map using the PWE simulator, taking into account the effects of rainfall. Specifically, we set the rain rate to $R = 25 \text{ mm/h}$. For simplicity, we utilize a unit-strength Gaussian beam as the radiating source at each base station, and the operating frequency is set to 4.9 GHz.

Algorithm 1 Dueling DDQN Multi-Step Learning for Trajectory Design in Rain Medium

```

1: Initialize: maximum number of episodes  $\bar{N}_{epi}$ , maximum number of steps per episode  $\bar{N}_{step}$ , update frequency  $N_{update}$ , reaching-destination toleration distance  $D_{tol}$ , initial exploration  $\epsilon_0$ , decaying rate  $\alpha$ , set  $\epsilon \leftarrow \epsilon_0$ .
2: Initialize: reaching-destination reward  $R_{des}$ , out-of-boundary penalty  $P_{ob}$ , SIR penalty weight, replay memory queue  $D$  with capacity  $C$  and minibatch size  $B$ .
3: Initialize: dueling DQN network with coefficients  $\theta$ , target network with coefficients  $\theta^- = \theta$ .
4: Initialize: SIR map in rain medium obtained from the PWE simulator.
5: for  $n_{epi} = 1, \dots, \bar{N}_{epi}$  do
6:   Initialize: a sliding window queue  $W$  with length  $N_1$ 
7:   Randomly set the starting point  $\mathbf{q}_I \in \mathcal{S}$ 
8:    $\mathbf{q}_0 \leftarrow \mathbf{q}_I, n \leftarrow 0$ 
9:   while  $\|\mathbf{q}_n - \mathbf{q}_F\| \geq D_{tol} \ \&\& \ \mathbf{q}_n \in \mathcal{S} \ \&\& \ n \leq \bar{N}_{step}$  do
10:    Generate a random number  $r \sim \mathcal{U}(0, 1)$ 
11:    if  $r < \epsilon$  then
12:      Action  $\vec{\mathbf{v}}_n \leftarrow \vec{\mathbf{v}}^k, k = \text{randi}(K)$ 
13:    else
14:      Action  $\vec{\mathbf{v}}_n \leftarrow \vec{\mathbf{v}}^k, k = \underset{k=1, \dots, K}{\text{argmax}} \hat{Q}(\mathbf{q}_n, \vec{\mathbf{v}}^{(k)}; \theta)$ 
15:    end if
16:     $\mathbf{q}_{n+1} \leftarrow \mathbf{q}_n + \vec{\mathbf{v}}_n$ ; sample maximum SIR at state  $\mathbf{q}_{n+1}$ ; set the reward  $R_n = 1 - \mu \text{SIR}$ ; store transition  $(\mathbf{q}_n, \vec{\mathbf{v}}_n, R_n, \mathbf{q}_{n+1})$  in the sliding window queue  $W$ 
17:    if  $n \geq N_1$  then
18:      Calculate the  $N_1$ -step accumulated return  $R_{(n-N_1):n}$  using (39).  $R_{n+i+1}, i \in 0, \dots, N_1 - 1$  are stored in sliding window,  $W$ 
19:      Store transition  $(\mathbf{q}_{n-N_1}, \vec{\mathbf{v}}_{n-N_1}, R_{(n-N_1):n}, \mathbf{q}_n)$  in the replay memory  $D$ 
20:    end if
21:    if  $|D| \geq B$  then
22:      Randomly sample a minibatch of  $N_1$ -step transition  $(\mathbf{q}_j, \vec{\mathbf{v}}_j, R_{j:(j+N_1)}, \mathbf{q}_{j+N_1})$  from  $D$ 
23:      for transition  $i$  in minibatch do
24:        if  $\|\mathbf{q}_{j+N_1} - \mathbf{q}_F\| \leq D_{tol}$  then
25:           $y_j = R_{j:(j+N_1)} + R_{des}$ 
26:        else if  $\mathbf{q}_{j+N_1} \notin \mathcal{S}$  then
27:           $y_j = R_{j:(j+N_1)} - P_{ob}$ 
28:        else
29:           $\vec{\mathbf{v}}^* = \underset{\vec{\mathbf{v}}' \in \mathcal{A}}{\text{argmax}} \hat{Q}(\mathbf{q}_{j+N_1}, \vec{\mathbf{v}}'; \theta)$ 
30:           $y_j = R_{j:(j+N_1)} + \gamma^{N_1} \hat{Q}(\mathbf{q}_{j+N_1}, \vec{\mathbf{v}}^*; \theta)$ 
31:        end if
32:      end for
33:      Perform a gradient descent step on  $\frac{1}{B} \sum_{j=0}^{B-1} (y_j - \hat{Q}(\mathbf{q}_j, \vec{\mathbf{v}}_j; \theta))^2$  w.r.t  $\theta$ 
34:      if  $\text{mod}(n, N_{update}) == 0$  then
35:         $\theta^- \leftarrow \theta$ 
36:      end if
37:    end if
38:     $n \leftarrow n + 1$ 
39:  end while
40: end for

```

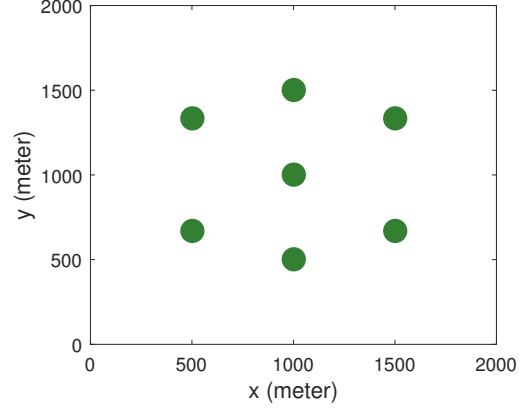


Fig. 6. Locations of the seven ground-based base stations (top view).

First, we use the PWE simulator to obtain the associated RSS in such an environment with varying weather conditions. Specifically, we examine a homogeneous atmosphere medium and a rain medium. The RSS distribution on a horizontal plane at a height of 100m in the airspace is compared in Fig. 7. The results show that the rainy environment has a significant impact on radio wave propagation. Therefore, accurate physics-based wave propagation models are necessary for UAV trajectory design under such weather conditions.

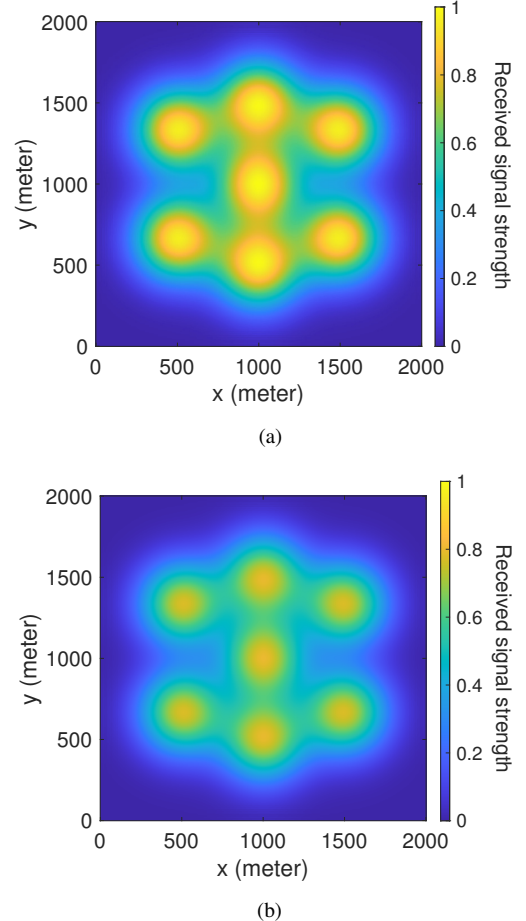


Fig. 7. Received signal strength in the airspace at a fixed height of 100m. (a) Homogeneous atmosphere medium. (b) Rain medium.

Second, dueling DDQN with multi-step learning is employed to optimize the UAV's trajectory. Specifically, we randomly select 200 initial starting points and set a fixed destination point. During the training process, the moving average return of the proposed algorithm is presented in Fig. 8. Additionally, the trajectories at different episodes of three

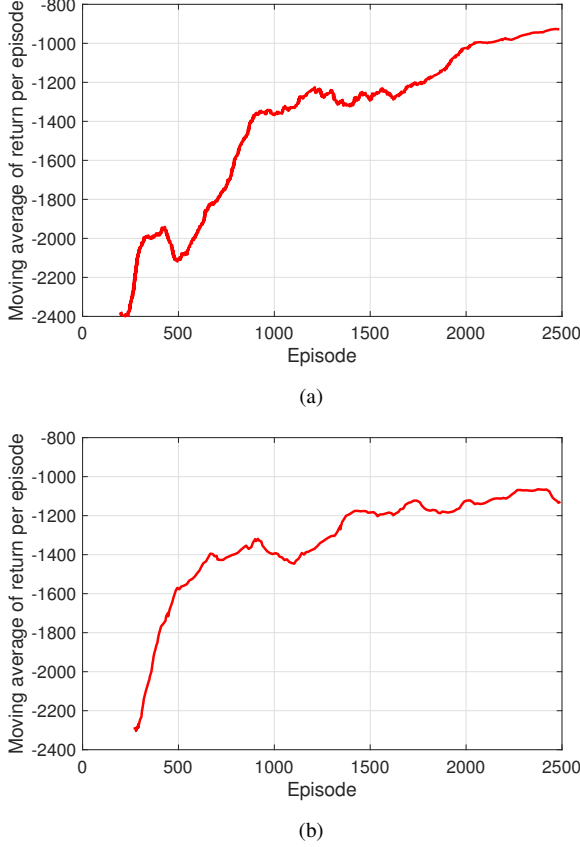


Fig. 8. Moving average return. (a) Homogeneous atmosphere medium. (b) Rain medium.

initial points with rain are presented in Fig. 9. The initial points are marked using red crosses, while the blue triangle indicates the final destination point. The minimum SIR is set as 10 dB in this section. It can be seen that the proposed algorithm converges as the number of episodes increases.

Upon convergence, the designed trajectories for the cellular-connected UAVs in the homogeneous atmosphere medium and the rain medium are plotted in Fig. 10. The parameters of the DRL algorithm in this section are summarized in Table I. Note that the same set of initial locations is used for the two cases with different mediums in Fig. 10. The proposed model is trained on an NVIDIA Geforce RTX 3090 GPU, and the execution time is 3480 s. The results show that the UAVs can avoid weak coverage regions and reach the destination point successfully. Additionally, the optimal trajectories for the same initial point differ between the two cases, indicating that the presence of rain has affected the trajectory.

Subsequently, we investigate the impact of two key parameters, namely, the weight factor, μ , and the rain rate, R , on the UAV trajectory design. In Fig. 11, UAV trajectories in cases with two different values of μ are plotted. The execution time of two cases with different μ is 3254 s and 3252 s, respectively.

TABLE I
PARAMETERS FOR THE PROPOSED TRAJECTORY DESIGN ALGORITHM

Parameter Meaning	Symbol	Value
maximum number of episodes	\bar{N}_{epi}	3000
maximum number of steps per episode	\bar{N}_{step}	300
update frequency	N_{update}	5
reaching-destination toleration distance	D_{tol}	10
initial exploration	ϵ_0	0.5
decaying rate	α	0.998
reaching-destination reward	R_{des}	2000
out-of-boundary penalty	P_{op}	10000
SIR penalty weight	μ	$\frac{10}{43}$
replay memory queue capacity	C	100000
minibatch size	B	16

As can be seen, when $\mu = 0.1$, the UAVs focus more on the flying time, and prefer to reach the destination point quickly. In contrast, when $\mu = 10$, the UAVs can better avoid the weak coverage regions. Thus, the choice of μ can be based on specific mission requirements.

To evaluate the impact of various rainy weather conditions on the UAV trajectory design, we conduct experiments with different rain rates. Specifically, UAV trajectories under different rain rates R are generated and compared in Fig. 12. We consider three distinct cases with rain rates set to $R = 25$ mm/h, $R = 50$ mm/h, and $R = 100$ mm/h, respectively. The execution time of three cases with different rain rates is 3194 s, 3208 s, and 3201 s, respectively. The experimental results demonstrate that our approach is capable of effectively handling different levels of rain, indicating its robustness and adaptability in practical scenarios with diverse weather conditions.

Additionally, the wind in a rainy medium can also affect wave propagation mainly due to the wind-induced movement of raindrops in a rainy medium introduces variations in the scattering environment [42]. A correction factor is employed by considering the individual drop-size dependent velocity resultant:

$$F(D) = \cos \left(\tan^{-1} \left(\frac{v_h}{v_v(D)} \right) \right) \quad (42)$$

where v_h is the horizontal wind speed. $v_v(D)$ indicates the drop diameter dependent terminal velocity and can be found

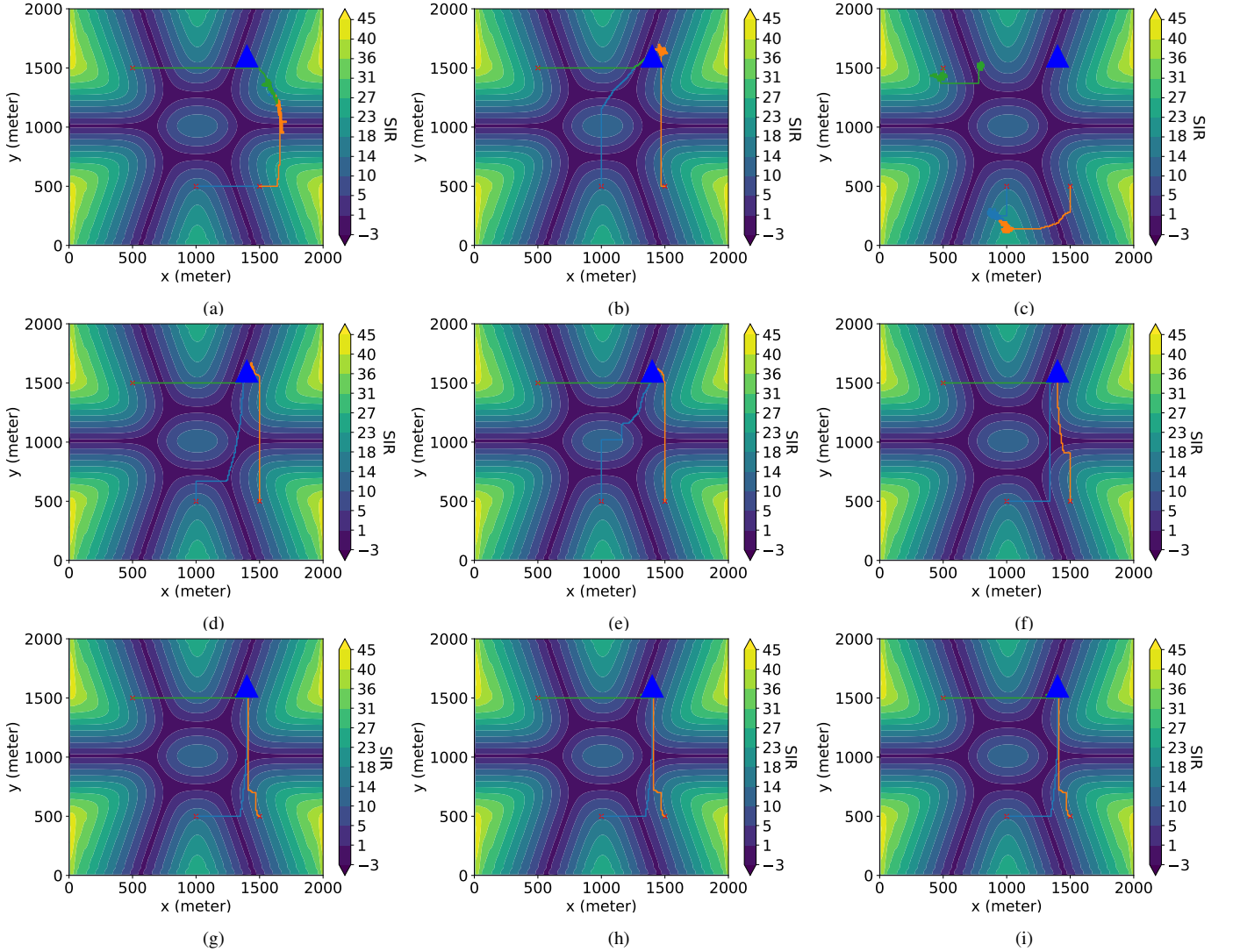


Fig. 9. Trajectories at different episodes during the training process. (a) 600 episodes. (b) 900 episodes. (c) 1200 episodes. (d) 1500 episodes. (e) 1800 episodes. (f) 2100 episodes. (g) 2400 episodes. (h) 2700 episodes. (i) 3000 episodes.

in [43]. Then, the effective permittivity in rain medium (19) can be obtained while considering the wind condition:

$$\begin{aligned}
 \varepsilon_{\text{eff}} = & \varepsilon_0 + \int_{D_{\min}}^{1.25} N(D) 4\pi \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega - 2\varepsilon_0} \left(\frac{D}{2}\right)^3 \\
 & \times \frac{3\varepsilon_0/F(D_i)}{3\varepsilon_0 - 4\pi \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega - 2\varepsilon_0} \left(\frac{D}{2}\right)^3} dD \\
 & + \frac{1}{3} \sum_{i=1}^3 \int_{1.25}^{D_{\max}} N(D) \frac{4\pi}{3} \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega + L_i(\varepsilon_\omega - \varepsilon_0)} \left(\frac{D}{2}\right)^3 \\
 & \times \frac{\varepsilon_0/F(D_i)}{\varepsilon_0 - L_i \frac{4\pi}{3} \frac{\varepsilon_0(\varepsilon_\omega - \varepsilon_0)}{\varepsilon_\omega + L_i(\varepsilon_\omega - \varepsilon_0)} \left(\frac{D}{2}\right)^3} dD.
 \end{aligned} \quad (43)$$

In Fig. 13, optimized trajectories considering about the wind condition are presented. The rain rates are all $R = 25$ mm/h. In the two cases considering the wind condition, the horizontal wind speed is set as $v_v(D) = 2$ m/s and $v_v(D) = 5$ m/s, respectively. The proposed model takes approximately 3192 s

to run in the case that accounts for the wind condition. It can be seen that the variations in the wind can affect radio wave propagation, leading to changes in the SIR distribution. Consequently, the UAVs' optimized trajectories need to be adjusted accordingly.

VII. CONCLUSION

This paper presents a novel physics-based trajectory design approach for cellular-connected UAVs in rainy environments. Compared to the previous optimal trajectory studies, the proposed approach takes into account detailed information about the environment's geometry and the impact of weather conditions on radio wave propagation and trajectory design. To formulate the trajectory design problem, we have defined the state, action space, and reward function, and a site-specific electromagnetic simulator and dueling DDQN with multi-step learning are utilized to solve this problem. The comparison of the optimal trajectory for cellular-connected UAVs in the

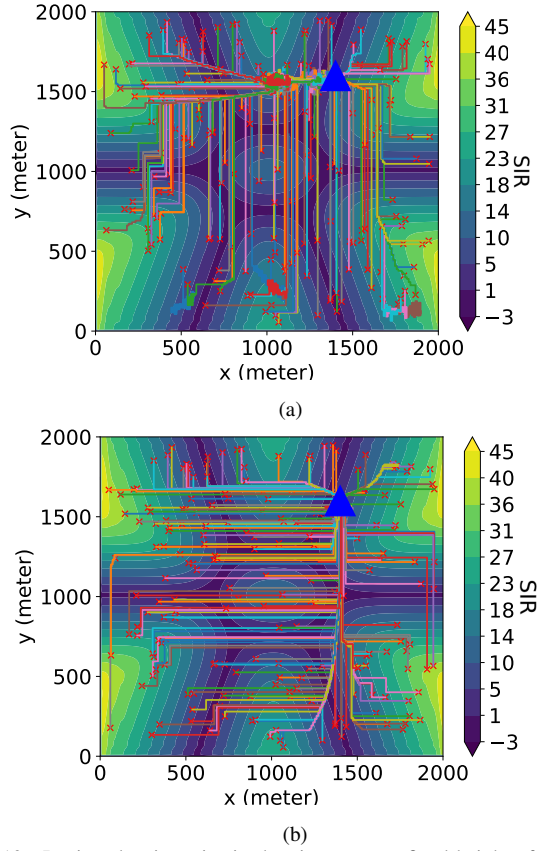


Fig. 10. Designed trajectories in the airspace at a fixed height of 100 m. (a) Homogeneous atmosphere medium. (b) Rain medium.

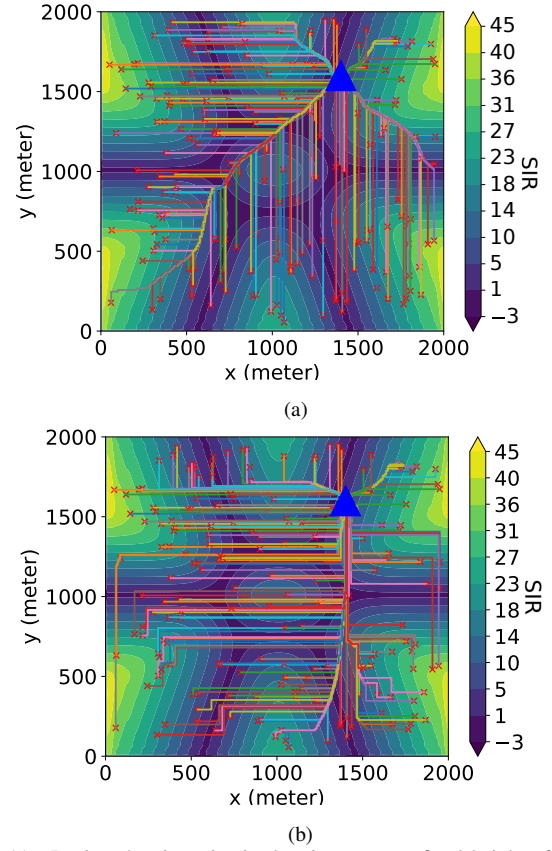


Fig. 11. Designed trajectories in the airspace at a fixed height of 100 m in rain medium with different μ . (a) $\mu = 0.1$. (b) $\mu = 10$.

atmosphere medium and rain medium demonstrates the necessity of our proposed approach. Additionally, we have provided a thorough study of varying weather conditions on trajectory design and investigated the impact of various parameters in the formulated problem. Overall, the proposed approach provides a valuable contribution to the field of UAV trajectory design in complex and challenging weather environments and has great potential for various UAV applications, such as weather monitoring and search and rescue operations.

REFERENCES

- [1] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.
- [2] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in uav communication networks," *IEEE Commun. Surv. & Tutor.*, vol. 18, no. 2, pp. 1123–1152, 2015.
- [3] Z. Yang, W. Xu, and M. Shikh-Bahaei, "Energy efficient UAV communication with energy harvesting," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1913–1927, 2019.
- [4] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A survey of channel modeling for UAV communications," *IEEE Commun. Surv. & Tutor.*, vol. 20, no. 4, pp. 2804–2821, 2018.
- [5] M. M. Azari, F. Rosas, K.-C. Chen, and S. Pollin, "Ultra reliable UAV communication using altitude and cooperation diversity," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 330–344, 2018.
- [6] R. Ding, F. Gao, and X. S. Shen, "3D UAV trajectory design and frequency band allocation for energy-efficient and fair communication: A deep reinforcement learning approach," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 12, pp. 7796–7809, 2020.
- [7] R. Amorim, H. Nguyen, P. Mogensen, I. Z. Kovács, J. Wigard, and T. B. Sørensen, "Radio channel modeling for UAV communication over cellular networks," *IEEE Wirel. Commun. Lett.*, vol. 6, no. 4, pp. 514–517, 2017.
- [8] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surv. & Tutor.*, vol. 21, no. 4, pp. 3417–3442, 2019.
- [9] S. Zhang, H. Zhang, B. Di, and L. Song, "Cellular UAV-to-X communications: Design and optimization for multi-UAV networks," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 2, pp. 1346–1359, 2019.
- [10] M. M. Azari, G. Geraci, A. Garcia-Rodriguez, and S. Pollin, "UAV-to-UAV communications in cellular networks," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 9, pp. 6130–6144, 2020.
- [11] C. Zhan and Y. Zeng, "Energy-efficient data uploading for cellular-connected UAV systems," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 11, pp. 7279–7292, 2020.
- [12] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, 2016.
- [13] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 3, pp. 2109–2121, 2018.
- [14] C. Zhan and Y. Zeng, "Energy minimization for cellular-connected UAV: From optimization to deep reinforcement learning," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 7, pp. 5541–5555, 2022.
- [15] Y.-J. Chen and D.-Y. Huang, "Joint trajectory design and BS association for cellular-connected UAV: An imitation-augmented deep reinforcement learning approach," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2843–2858, 2021.
- [16] S. Hu, X. Yuan, W. Ni, and X. Wang, "Trajectory planning of cellular-connected UAV for communication-assisted radar sensing," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 6385–6396, 2022.
- [17] S. Zhang and R. Zhang, "Trajectory optimization for cellular-connected

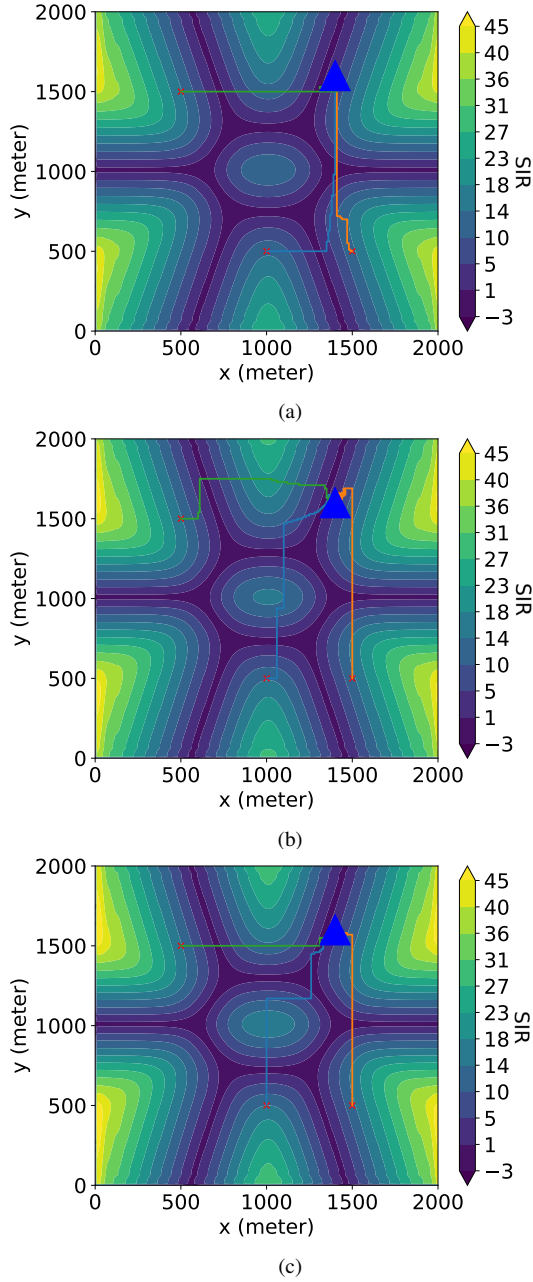


Fig. 12. Designed trajectories in the airspace at a fixed height of 100 m with different rain rates. (a) $R = 25$ mm/h. (b) $R = 50$ mm/h. (c) $R = 100$ mm/h.

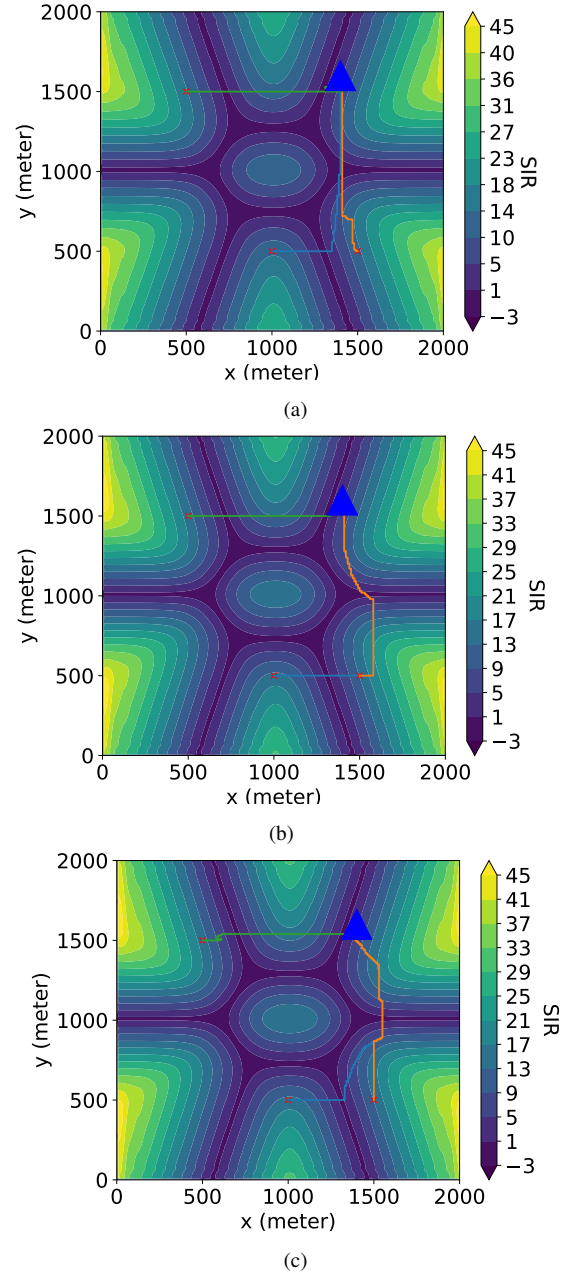


Fig. 13. Designed trajectories in the airspace at a fixed height of 100 m. (a) Rain medium without wind condition. (b) Rain medium with wind condition (wind speed: 2 m/s). (c) Rain medium with wind condition (wind speed: 5 m/s).

uav under outage duration constraint,” *J. Commun. Inf. Netw.*, vol. 4, no. 4, pp. 55–71, 2019.

- [18] Y. Zeng, X. Xu, and R. Zhang, “Trajectory design for completion time minimization in UAV-enabled multicasting,” *IEEE Trans. Wirel. Commun.*, vol. 17, no. 4, pp. 2233–2246, 2018.
- [19] S. Zhang, Y. Zeng, and R. Zhang, “Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective,” *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2580–2604, 2019.
- [20] Y. Zeng and R. Zhang, “Energy-efficient UAV communication with trajectory optimization,” *IEEE Trans. Wirel. Commun.*, vol. 16, no. 6, pp. 3747–3760, 2017.
- [21] S. Zhang and R. Zhang, “Radio map-based 3D path planning for cellular-connected UAV,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 3, pp. 1975–1989, 2020.
- [22] U. Challita, W. Saad, and C. Bettstetter, “Deep reinforcement learning for interference-aware path planning of cellular-connected UAVs,” *2018 IEEE Int. Conf. Communi. (ICC)*, pp. 1–7, 2018.
- [23] U. Challita, A. Ferdowsi, M. Chen, and W. Saad, “Machine learning for

wireless connectivity and security of cellular-connected UAVs,” *IEEE Wirel. Commun.*, vol. 26, no. 1, pp. 28–35, 2019.

- [24] B. Khamidehi and E. S. Sousa, “A double Q-learning approach for navigation of aerial vehicles with connectivity constraint,” *2020 IEEE Int. Conf. Communi. (ICC)*, pp. 1–6, 2020.
- [25] Y. Zeng, X. Xu, S. Jin, and R. Zhang, “Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 7, pp. 4205–4220, 2021.
- [26] U. Challita, W. Saad, and C. Bettstetter, “Interference management for cellular-connected UAVs: A deep reinforcement learning approach,” *IEEE Trans. Wirel. Commun.*, vol. 18, no. 4, pp. 2125–2140, 2019.
- [27] M. Y. Arafat, M. M. Alam, and S. Moh, “Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges,” *Drones*, vol. 7, no. 2, p. 89, 2023.
- [28] A. Abdulrahman, T. Rahman, S. Rahim, and M. U. Islam, “Empirically

- derived path reduction factor for terrestrial microwave links operating at 15 GHz in peninsula malaysia,” *J. Electromagn. Waves Appl. J.*, vol. 25, no. 1, pp. 23–37, 2011.
- [29] A. Abdulrahman, T. A. Rahman, S. K. A. Rahim, M. R. Islam, and M. Abdulrahman, “Rain attenuation predictions on terrestrial radio links: differential equations approach,” *Trans. Emerg. Telecommun. Technol.*, vol. 23, no. 3, pp. 293–301, 2012.
- [30] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*. London, U.K.: Inst. Elect. Eng., 2000.
- [31] N. Sheng, X.-M. Zhong, Q. Zhang, and C. Liao, “Study of parabolic equation method for millimeter-wave attenuation in complex meteorological environments,” *Prog. Electromagn. Res. M*, vol. 48, pp. 173–181, 2016.
- [32] N. Sheng, C. Liao, W. Lin, Q. Zhang, and R. Bai, “Modeling of millimeter-wave propagation in rain based on parabolic equation method,” *IEEE Antennas and Wireless Propag. Lett.*, vol. 13, pp. 3–6, 2013.
- [33] Z. He, T. Su, H.-C. Yin, and R.-S. Chen, “Wave propagation modeling of tunnels in complex meteorological environments with parabolic equation,” *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 6629–6634, 2018.
- [34] H. Qin and X. Zhang, “Efficient modeling of radio wave propagation in tunnels for 5G and beyond using a split-step parabolic equation method,” *IEEE General Ass. and Sci. Symp. of the Inte. Union of Radio Science (URSI GASS)*, pp. 1–3, 2021.
- [35] S. Gong and J. Huang, “Accurate analytical model of equivalent dielectric constant for rain medium,” *J. Electromagn. Waves Appl. J.*, vol. 20, no. 13, pp. 1775–1783, 2006.
- [36] H. J. Liebe, T. Manabe, and G. A. Hufford, “Millimeter-wave attenuation and delay rates due to fog/cloud conditions,” *IEEE Trans. Antennas Propag.*, vol. 37, no. 12, pp. 1617–1612, 1989.
- [37] *Specific Attenuation Model for Rain for Use in Prediction Methods*. ITU-R Recommendation P.838-3, ITU-R, Geneva, Switzerland, 2005.
- [38] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [39] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, “Human-level control through deep reinforcement learning,” *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [40] M. Hessel, J. Modayil, H. Van Hasselt, T. Schaul, G. Ostrovski, W. Dabney, D. Horgan, B. Piot, M. Azar, and D. Silver, “Rainbow: Combining improvements in deep reinforcement learning,” in *Proc. of the AAAI conf. on artif. intel.*, vol. 32, no. 1, 2018.
- [41] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas, “Dueling network architectures for deep reinforcement learning,” in *Inte. conf. on machine learning*. PMLR, 2016, pp. 1995–2003.
- [42] W. Asen and T. Tjelta, “A novel method for predicting site dependent specific rain attenuation of millimeter radio waves,” *IEEE Trans. Antennas Propag.*, vol. 51, no. 10, pp. 2987–2999, 2003.
- [43] G. Brussaard and P. A. Watson, *Atmospheric modelling and millimetre wave propagation*. Springer Science & Business Media, 1994.