2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC),
18-22 September 2022, Portsmouth, Virginia, USA
DOI:10.1109/DASC55683.2022.9925817

Interference Mitigation for 5G-Connected UAV using Deep Q-Learning Framework

Anirudh Warrier
School of Aerospace, Transport and Manufacturing(SATM)
Cranfield University
Bedford, United Kingdom
anirudh.warrier@cranfield.ac.uk

Saba Al-Rubaye
School of Aerospace, Transport and Manufacturing(SATM)
Cranfield University
Bedford, United Kingdom
s.alrubaye@cranfield.ac.uk

Dimitrios Panagiotakopoulos
School of Aerospace, Transport and Manufacturing(SATM)
Cranfield University
Bedford, United Kingdom
d.panagiotakopoulos@cranfield.ac.uk

Gokhan Inalhan
School of Aerospace, Transport and Manufacturing(SATM)
Cranfield University
Bedford, United Kingdom
inalhan@cranfield.ac.uk

Antonios Tsourdos
School of Aerospace, Transport and Manufacturing(SATM)
Cranfield University
Bedford, United Kingdom
a.tsourdos@cranfield.ac.uk

***Abstract*—To boost large-scale deployment of unmanned aerial vehicles (UAVs) in the future, a new wireless communication paradigm namely cellular-connected UAVs has recently received an upsurge of interest in both academia and industry. Fifth generation (5G) networks are expected to support this large-scale deployment with high reliability and low latency. Due to the high mobility, speed, and altitude of the UAVs there are numerous challenges that hinder its integration with the 5G architecture. Interference is one of the major roadblocks to ensuring the efficient co-existence between UAVs and terrestrial users in 5G networks. Conventional interference mitigation schemes for terrestrial networks are insufficient to deal with the more severe air-ground interference, which thus motivates this paper to propose a new algorithm to mitigate interference. A deep Q-learning (DQL) based algorithm is developed to mitigate interference intelligently through power control. The proposed algorithm formulates a non-convex optimization problem to maximize the Signal to Interference and Noise Ratio (SINR) and solves it using DQL. Its performance is measured as effective SINR against the complement cumulative distribution function. Further, it is compared with an adaptive link technique: Fixed Power Allocation (FPA), a standard power control scheme and tabular Q-learning(TQL). It is seen that the FPA has the worst performance while the TQL performs slightly better. This is since power control and interference coordination are introduced but not as effectively in the TQL method. It is observed that DQL algorithm outperforms the TQL implementation. To solve the severe air-ground interference experienced by the UAVs in 5G networks, this paper proposes a DQL algorithm. The algorithm effectively mitigates interference by optimizing SINR of the air-ground link and outperforms the existing methods. This paper therefore, proposes an effective algorithm to resolve the interference challenge in air-ground links for 5G-connected UAVs.**

***Index Terms*—Fifth-generation(5G), interference, deep Q-learning, unmanned aerial vehicles (UAVs)**

Engineering and Physical Sciences Research Council(EPSRC), Satellite Applications Catapult

I. Introduction

Unmanned aerial vehicles (UAVs) have been gaining increasing popularity in recent years, especially as an enabler for a plethora of new applications, such as cargo delivery, surveillance and inspection, aerial photography, among others [1]. Advances in communication technology for instance miniaturization of hardware, have enabled UAVs to serve as communication interfaces in the sky such as base stations (BSs), to enhance the communication services for the terrestrial or aerial users in demand.

To support the deployment of UAVs on a large scale, they must be integrated into the future cellular network i.e., the fifth generation (5G) as BSs to support terrestrial BSs [2]. Compared to the existing UAV-ground communications, which is limited to within visual line-of-sight (VLoS) range, cellular-connected UAVs can only be enabled by the beyond visual LoS (BVLoS) communications. This leads to significant performance enhancement in terms of reliability, coverage, security and throughput [3].

Despite advantages, integrating UAVs into 5G networks faces challenges [4] [5]. In particular, how to mitigate the severe aerial-ground interference is deemed as a major question in realizing cellular-connected UAVs. Compared to terrestrial wireless channels that generally suffer from severe path-loss, shadowing and multi-path fading [6], the high altitude of UAVs leads to LoS-dominant channels with ground BSs. Due to these LoS links from non-associated BSs, the UAV may suffer severe downlink interference, which would significantly degrade the communication performance of UAVs.

A. Interference Mitigation Challenges

As mentioned above, one of the major challenges in the functioning of UAVs as part of a cellular network lies in severe air-ground interference that it faces. Compared to traditional ground users, interference in UAVs is aggravated by the LoS-dominated UAV-BS channels owing to the altitude of the UAV. For downlink communication from BS to UAV, each UAV may receive severe interference from a large number of neighboring BSs that are not associated with it, due to strong LoS-dominated channels. As a result, it is expected that a UAV in general would have a poor downlink performance. On the other hand, in the uplink communication from UAV to BS, the UAV could also pose strong interference to many adjacent but non-associated BSs and result in a new exposed BS interference issue. Thus, devising an effective interference mitigation technique by taking into account the unique UAV-BS channel and interference characteristics is crucial to cellular-connected UAVs. One simple method of mitigating this interference is efficient allocation of resource blocks(RBs), but for severe air-ground interference this cannot be used. This is because the number of RBs available to the UAVs in this case would be highly limited or even zero with a high probability of outage, due to the dense frequency reuse for terrestrial UEs in today's cellular network and their current number significantly surpasses that of UAVs. Furthermore, existing terrestrial mitigation techniques have their limitations in dealing with the more severe air-ground interference.

The aim of this paper is to develop a unique method of power control to mitigate interference in 5G networks that cope with the unique challenges posed by the use of UAVs. A deep Q-learning algorithm is proposed to mitigate interference via power control. The following explains why the proposed DQL is a good solution to the problem:

- The proposed solution does not require the channel state information (CSI) (as opposed to standard link adaptation techniques), to find the optimal Signal-to-Interference and Noise-Ratio(SINR).
- It also reduces the need for UAV feedback to the BS. In existing methods [7], the UE or the UAV has to report its CSI which is a vector of length equal to number of antenna elements. In the proposed method, the UAVs send only their received SINR and co-ordinates, while the agent manages the power control and consequently the interference coordination commands.
- The proposed solution provides explicit power control and interference mitigation commands sent by the UAV to the serving and interfering BSs as opposed to current industry standards which only require the serving BS to send power control commands to the UAV.

B. Contributions

This paper therefore, makes the following contributions:

- Formulate the power control, and interference problem in the downlink direction as an optimization problem that maximizes the UAV's received SINR.
- To create a deep reinforcement learning based solution where multiple actions can be taken at once using information in a data set and change the power values to achieve optimal SINR.

The paper is structured as follows, Section II discusses the existing works and identification of the gaps and the challenges associated with present interference mitigation techniques. 5G and UAV functionality is discussed in detail in Section III. A brief introduction on deep reinforcement learning and existing techniques of interference mitigation against which the proposed solution is compared are presented, further the proposed algorithm is also presented in Section IV. Network model, system model, and the channel models are discussed along with the simulation setup and results in Section V. The work is concluded and the impact of the results discussed in Section VI.

II. Background and Related Works

Although various interference mitigation techniques have been studied in the literature some of which were applied to the terrestrial networks such as inter-cell interference coordination (ICIC) [8] [9], coordinated multi-point (CoMP) transmission [10] [11] they may be inadequate to deal with the new and more severe interference issue brought by UAVs, owing to their unique LoS-dominant air-ground channels. The industry standards adopted the method of almost blank sub-frame (ABS) to resolve the co-channel inter-cell interference problem in LTE where two BSs interfere with one another. ABS works well in fixed beam antenna patterns, the dynamic nature of UAVs reduces the usefulness of ABS for UAV networks. In [12], the interference characteristics of directional UAV networks is characterized based on the stochastic geometry, where each UAV is equipped with a directional antenna and is placed in three dimensional (3D) locations. In particular, the 3D location of UAVs is assumed to be uniformly distributed in a certain volume, which is modeled by Poisson point process. As discussed earlier, UAVs in cellular network as new aerial users is a promising solution to meet their ever-increasing communication demands, but owing to the high UAV altitude, the channels are dominated by the strong LoS links. A UAV can communicate with a large number of base stations at the same time, leading to a higher probability of interference as compared to terrestrial users. However, on the other hand, severe interference may be generated from the LoS links, which renders the interference management with coexisting terrestrial and aerial users a more challenging problem to solve. In [13], the authors propose a new cooperative interference cancellation strategy for the multi-beam UAV uplink communication, which aims to eliminate the co-channel interference at each of the occupied BSs and maximize the sum-rate to the available BSs. Over the last few years, the use of deep learning in wireless communications was studied in certain literature [14]. Specifically, [15] uses deep reinforcement learning to perform power control for mmWave and this was designed as an alternative to beamforming in improving the non-line of sight (NLOS) transmission performance. The power allocation problem to maximize the sum-rate of UEs under the constraints of transmission power and quality targets was solved using deep reinforcement learning. In their solution, the authors use a convolutional neural network to estimate the Q-function of the deep reinforcement learning problem. [16] defines a policy that maximizes the successful transmissions in a dynamic co-related multichannel access environment as obtained using deep Q -learning. In [17], the authors jointly optimized beamforming, power control, and interference coordination in a 5G wireless network to enhance the communication performance to end users. They developed Q-learning algorithm to maximize the downlink SINR in a multi-access OFDM cellular network from a multi-antenna base station to single-antenna UEs.

It is seen that, the solutions provided to mitigate interference for UAVs are developed on existing interference cancellation, sum-rate maximization and beamforming techniques but these only work for certain use cases. With UAVs touted as the solution to wide range of challenges and applications, it is of importance that in future 5G networks is able to accommodate UAVs operating at high altitude and speeds with minimum possible interference.

III. 5G-UAV Functionality

Using 5G networks for UAVs is an opportunity to provide stable connectivity, while reducing scale, weight and power usage costs and specifications. 5G is expected to provide a wide variety of wireless services across multiple access channels and multi-layer networks. To that end, it uses a smarter Radio Access Networks (RANs) architecture, that is not limited by the proximity of the base station or complex infrastructure. Some of the technology enablers that support 5G networks and impacts the UAVs are, 5G spectrum and frequency, beamforming, multi access edge computing, software defined networks and network function virtualization and network slicing [18].

According to 3GPP standard [19], the new 5G Network Core uses a cloud Service-Based Architecture (SBA), as Fig. 1 depicts, that covers all 5G functions and interactions, such as authentication, protection, session management and end-user traffic aggregation. It emphasizes virtualised functions deployed using the Multi-access Edge Computing(MEC) infrastructure, as an integral design principle. As an advancement of cloud computing, MEC is a crucial element of the 5G infrastructure that takes applications from centralized data centres to the edge network, which means closer to the end user, offering advantages such as low latency, high bandwidth, and real-time access to RAN information. It also guarantees very low-latency communication for the C2 link.

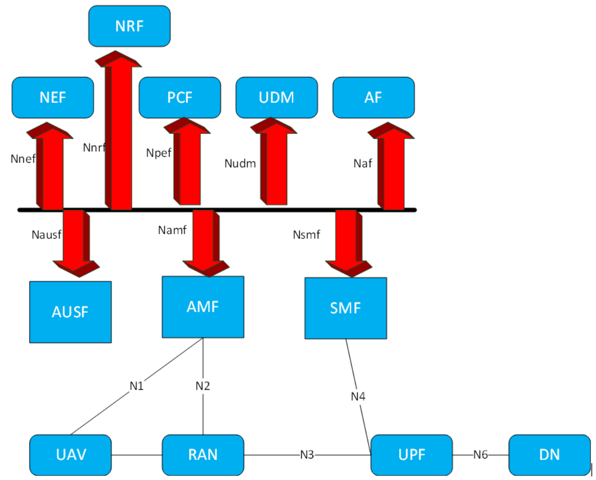

Fig. 1. 5G System Architecture

The system considered in this paper, consists of primarily of the following components:

1) UAV: The UAV consists of the following sub-components as shown in Fig. 2 which serve different functions to enable different functionalities for the UAV:

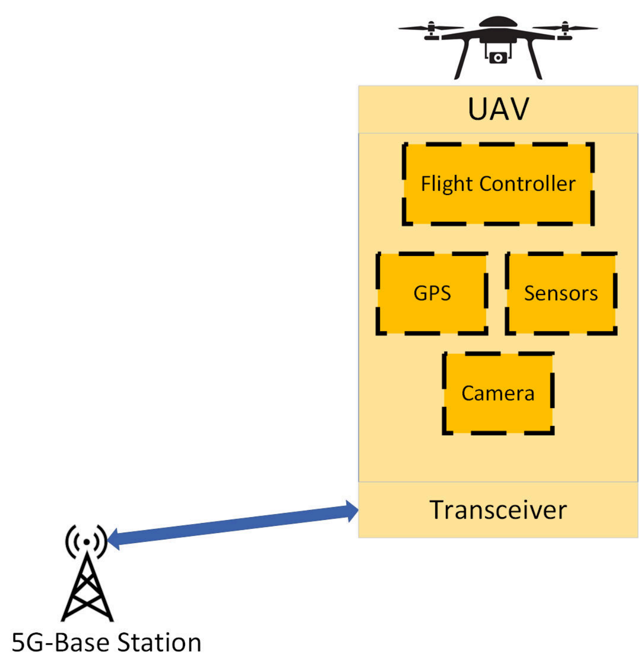

Fig. 2. UAV functionality

- Flight Controller: The flight controller is responsible for mainly three different functions, viz. (i) sensing

- sensors give the flight controller information like its height, orientation, and speed. Common sensors include an Inertial Measurement Unit (IMU) for determining the angular speed and acceleration, a barometer for the height, and distance sensors for detecting obstacles, (ii) controlling - UAVs can rotate and accelerate by creating speed differences between each of its four motors. The flight controller uses the data gathered by the sensors to calculate the desired speed for each of the four motors, (iii) communicating - flight controllers need to communicate with other computer systems about its flight destination, UAV health and other mission critical data.

- GPS: A GPS module allows UAVs to know their location relative to a network of orbiting satellites. Connecting to signals from these satellites allows the UAV to perform functions such as position hold, autonomous flight, return to home, and way-point navigation.

- Sensors: Sensors for UAVs are used for surveying, mapping, and inspections, to support a wide range of application in industries such as mining, construction, energy, environmental management, agriculture, infrastructure, and waste management. Sensor advances are capable of improving multi-functionality to allow for a wider range of applications.

Additionally the UAV also consists of a transceiver, that sends and receives data from the ground control station (GCS) which in this case is a 5G base station. In the proposed algorithm, the UAV sends the received SINR values along with its co-ordinates to the base station, based on which the interference is mitigated. The SINR values are reported from the flight controller unit, while the co-ordinates are reported for the GPS module.

2) Ground Control Station(GCS): Ground control stations are mainly responsible for the following functions- attitude control of the UAV, display and control of pay-load data, mission planning, UAV position monitoring, map display of routes,navigation and target positioning, and communication links with other subsystems.

In the system considered, power control and interference mitigation on the signal from the non-serving BS is performed at a central location. The decisions are computed at a central location, which is located at in a cloud architecture. The measurements from the UAVs are relayed to the central location over the backhaul as shown in Fig. 3 which is further linked to the 5G core network.

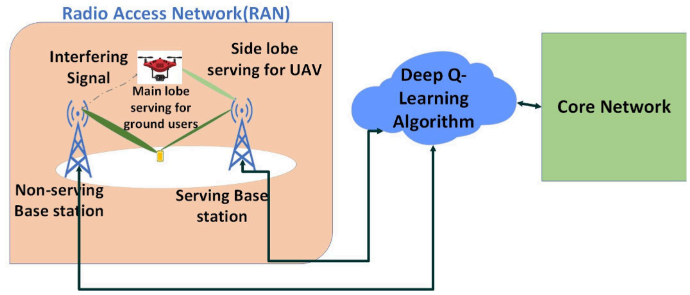

Fig. 3. Base station functionality

Therefore, by modifying functionalities at the BS end, the deep reinforcement learning algorithm is implemented to mitigate interference by improving SINR of the UAV.

IV. Deep Reinforcement Learning

Deep Reinforcement Learning (DRL) is a machine learning technique in which an *agent* is enabled to discover a certain *action* it should take to maximize its expected future *reward* in an interactive environment. This interaction is shown Fig. 4. DRL exploits the capability of deep neural networks to learn better representations and operate as a universal function approximator.

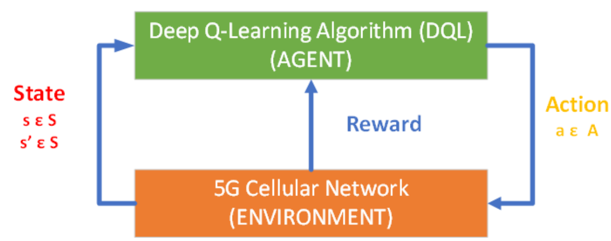

Fig. 4. Agent-environment interaction

A. RL learning elements

Reinforcement learning has several elements [20]. These elements interact together, and are as follows:

- Observations: Observations are continuous measures of the properties of the environment and are written as a p -ary vector $\mathcal{O} \in \mathbb{R}^p$ where p is the number of properties observed.
- States: The state $s_t \in \mathcal{S}$ is the discretization of the observations at time step t. Often, states are also used to mean observations.

- Actions: An action $a_t \in \mathcal{A}$ is one of the valid choices that the agent can make at time step t. The action changes the state of the environment from the current state s to the target state s'.

- Policy: A policy $\pi(\cdot)$ is a mapping between the state of the environment and the action to be taken by the agent.

A stochastic policy is $\pi(a|s) : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$.

- Rewards: The reward signal $r_{s,s',a}[t;q]$ is obtained after the agent takes an action a when it is in state s at time step t and moves to the next state s'.
- State-action value function: The state-action value function under a given policy π is denoted $Q_\pi(s, a)$. It is the expected discounted reward when starting in state s and selecting an action a under the policy π.

These elements work together and their relationship is governed by the objective to maximize the future discounted reward for every action chosen by the agent, which causes the environment to transition to a new state. The policy dictates the relationship between the agent and the state. The value of the expected discounted reward is learned through the training phase.

B. Existing Methods

In this paper, the proposed algorithm is compared with some industry standards, which are discussed as follows:

1) Fixed Power Control (FPA): The fixed power allocation (FPA) power control is used as a baseline algorithm that sets the transmit signal power at a specific value. No interference coordination is implemented in FPA. Total transmit power is divided equally among all the Physical Resource Blocks (PRBs) and is therefore constant. In this algorithm, the BS fixes its transmit power and only changes the modulation and code schemes of the transmission. This is known as the "link adaptation." Link adaptation takes place based on the measurement reports sent by the UAV back to the BS (i.e., the SINR and received power). Since the BS transmit power is fixed, the link adaptation takes place based on either periodic or aperiodic measurement feedback from the UAV to the serving BS. This results in an improved effective SINR and a reduction in the packet error rate. There is no measurement sent to the interfering BS based on FPA.

2) Tabular RL: The tabular setting of Q -learning (or "vanilla" Q -learning) is used to implement the algorithm for interference mitigation. In a tabular setting, the state-action value function $Q_\pi(s_t, a_t)$ is represented by a table $Q \in \mathbb{R}^{|S| \times |A|}$. The learning rate of the Q-learning update, defines how the experience update takes place with respect to the previous experience. Computationally, the tabular setting suits problems with small state spaces where maintaining a Q-table is possible.

C. Proposed Solution

Algorithm 1 is proposed, which is a DRL-based approach. This algorithm performs power control without the UAV sending explicit power control commands. This Deep Q-Learning (DQL) may provide a lower computational overhead compared to the tabular Q-learning depending on the number of states and the depth of the deep Q-network [17].

Algorithm 1 Deep Q-Learning Algorithm for Interference Mitigation

Input: Downlink SINR measured and reported by the UAVs
Output: Sequence of power control and interference coordination commands to optimize the SINR

```
Initialize time,states,actions,and replay buffer D
repeat
    repeat
        t := t + 1.
        Observe Current state s_t
        ε := max(ε · d, ε_min)
        Sample r ~ Uniform(0,1)
        if r ≤ ε then
            Select an action a_t ∈ A at random
        else
            An action a_t = arg max_{a'} Q_π(s_t, a'; θ_t)
        end if
        Compute γ^l_eff[t] and r_{s,s',a}[t;q]
        if γ^l_eff[t] < γ_min then
            r_{s,s',a}[t;q] := r_min
            Abort episode
        end if
        Observe next state s'
        Store experience e[t](s_t, a_t, r_{s,s',a'}, s') in D
        Minibatch sample from D for experience
                    e_j ≜ (s_t, a_t, r_j, s_{j+1})
        Set y_i := r_j + y max_{a'} Q_π(s_{j+1}, a'; θ_t)
        Perform SGD on (y_i − Q_π(s_j, a_j; θ_t)^2), find θ*
        Update θ_t := θ* in the DQN and record loss L_t
        s_t := s'
    until t ≥ T
until convergence or aborted
if γ^l_eff[t] ≥ γ_target then
    r_{s,s',a}[t;q] := r_{s,s',a}[t;q] + r_max
end
```

The main steps of Algorithm 1 are as follows:

1) Select an optimization action at a time step t.
2) Select a power control and interference coordination action.
3) Assess the impact on effective SINR $\gamma^l_{eff}[t]$.
4) Reward the action taken based on the impact on $\gamma^l_{eff}[t]$ and its distance from γ_{target} or γ_{min}, i.e. higher reward for being closer to γ_{target} and lower for being closer to γ_{min}.

5) Train the DQN based on the outcomes.

The notations and abbreviations used in the algorithm are described in Table I

TABLE I
LIST OF NOTATIONS

Notation	Description
t	Time Sample
s_t	State Variable
ϵ	Exploration rate
d	Exploration rate decay
a_t	Action Variable
Q_π	State Action Value Function
γ	SINR
γ_{eff}	Effective SINR
γ_{target}	Target SINR
γ_{min}	Minimum SINR
γ_{thresh}	Threshold SINR
SGD	Stochastic Gradient Descent

V. SIMULATION SET-UP AND RESULTS

A. Network Model

A downlink cellular network of L BSs is considered. This network is comprised of a serving BS and at least one interfering BS. A downlink scenario, where a BS is transmitting to the UAV is adopted. The BSs have an inter-site distance of R and the UAVs are randomly scattered. The association between the UAVs and their serving BS is based on the distance between them. A user is served by one BS maximum. The cell radius is $r > R/2$ to allow overlapping of coverage.

B. System Model

Considering the above explained network model, and adopting a multi-antenna setup where each BS employs a uniform linear array of M antennas and the UAVs have single antennas, the received signal at the UAV from the ℓ-th BS can be written as :

$$y_l = h_{l,l}^* f_l x_l + \sum_{b \neq l} h_{l,b}^* f_b x_b + n_l \tag{1}$$

where x_{l}, $x_{\text{b}} \in \mathbb{C}$ are transmitted signals from the ℓ-th and ℓ-th BSs, and it satisfies the power constraint

$$\mathbb{E}[|x_l|^2] = P_{\text{TX,l}} \tag{2}$$

The $M \times 1$ vectors $h_{\text{l,l}}, h_{\text{l,b}} \in \mathbb{C}^{M \times 1}$ are the channel vectors connecting the UAV at the ℓth BS with the ℓth and ℓth BSs respectively. Finally, $n_{\text{l}} \sim \text{Normal}(0, \sigma_n^2)$ is the received noise at the user sampled from a complex Normal distribution with zero-mean and variance σ_{n}^2.

The first term in (1) represents the desired received signal. The interference received by the UAV from the non-associated BSs, is depicted by the second term in (1).

Every BS ℓ is assumed to have a transmit power $P_{\text{TX,l}} \in \mathcal{P}$, where $\mathcal{P}$ is the set of possible transmit powers. This set is defined such that the possible transmit powers is a power offset (above or below the current power level).

C. Channel Model

A narrow-band geometric channel is adopted for this algorithm. With this geometric model, the downlink channel from a BS ℓ to the UAV in BS ℓ can be written as

$$h_{\text{l,b}} = \frac{\sqrt{M}}{\rho_{\text{l,b}}} \sum_{p=1}^{N_{l,b}^p} \alpha_{l,b}^p a^*(\theta_{l,b}^p) \tag{3}$$

where,

$\alpha_{l,b}^p =$ complex path gain of the p-th path

$\theta_{l,b}^p =$ angle of departure(AoD) of the p-th path

$a(\theta_{l,b}^p) =$ array response vector associated with AoD

$N_{l,b}^p =$ No. of channel paths

This model accounts for both LOS and NLOS scenarios. For the LOS case, $N_{l,b}^p = 1$ is assumed

The received downlink power as measured by the UAV over a set of physical resource blocks (PRBs) at a given time t as

$$P_{\text{UAV}}^{\text{l,b}}[t] = P_{\text{TX},b}[t] \left| h_{\text{l,b}}^*[t] f_b[t] \right|^2 \tag{4}$$

where,

$P_{\text{UAV}}^{\text{l,b}}[t]$- Received downlink power

$P_{\text{TX},b}[t]$- Transmit power from BS ℓ

Next the received SINR for the UAV served in BS ℓ at time t is computed as follows,

$$\gamma^\ell[t] = \frac{P_{\text{TX},\ell}[t] |h_{\ell,\ell}^*[t] f_\ell[t]|^2}{\sigma_n^2 + \sum_{b \neq \ell} P_{\text{TX},b}[t] |h_{\ell,b}^*[t] f_b[t]|^2} \tag{5}$$

This is the received SINR that will be optimized.

D. Simulation Setup

The network, system, and channel models are described in earlier sections. The users are moving at a speed v with both log-normal shadow fading and small-scale fading. The cell radius is r and the inter-site distance R=1.5r. The UAVs experience a probability of line of sight of p_{LOS}. The rest of the parameters are shown in Table II. The target effective SINRs are set as:

$$\begin{aligned} \gamma_{\text{target}} &:= 3 \text{ dB}, \\ \gamma_{\text{target}}^{\text{thresh}} &:= \gamma_0^{\text{thresh}} + 10 \log M \text{ dB} \end{aligned} \tag{6}$$

where γ_0^{thresh} is a constant threshold. A minimum SINR of -3 dB below which the episode is declared aborted and the session is unable to continue, is set.

The hyper parameters required to tune the RL-based model are shown in Table II. Further, we run Algorithm 1 on the cellular network with its parameters in Table III.

TABLE II
REINFORCEMENT LEARNING HYPER PARAMETERS

Parameter	Value
Discount Factor γ	0.995
Initial exploration rate ϵ	1.000
Number of States $\mathcal{S}$	8
Deep Q-Network width H	24
Exploration rate decay d	0.995
Minimum exploration rate $\epsilon_{min}, \epsilon_{min}^{thresh}$	(0.15,0.10)
Number of Actions $\mathcal{A}$	16
Deep Q Network Depth	2

The simulated states $\mathcal{S}$ are setup as:

$$(s_t^0, s_t^1) := \mathrm{UAV}_\ell(x[t], y[t]), (s_t^2, s_t^3) := \mathrm{UAV}_b(x[t], y[t]),$$
$$s_t^4 := P_{\mathrm{TX},\ell}[t], \quad s_t^5 := P_{\mathrm{TX},b}[t],$$
$$s_t^6 := \mathbf{f}_n^\ell[t], \quad s_t^7 := \mathbf{f}_n^b[t],$$

where (x,y) are the Cartesian co-ordinates (i.e longitude and latitude) of the given UAV.

TABLE III
RADIO ENVIRONMENT PARAMETERS

Parameter	Value
BS maximum transmit power $P_{BS}^{\max}$	46 dBm
Cellular Geometry	Hexagonal
Antenna Gain(TX,thresh)	(11,3) dBi
Probability of LOS ($P_{\mathrm{LOS}}, P_{\mathrm{LOS}}^{\mathrm{thresh}}$)	(0.9,0.8)
Downlink Frequency	4.7 GHz
Cell Radius r	150m
UAV Antenna Gain	0 dBi
Inter-site distance R	225m
Number of Multipaths N_p	15
Average UAV Speed v	20m/s
Frame Duration	20ms

E. Results

Fig. 5 shows the Complementary Cumulative Distribution Function (CCDF) of the effective SINR γ_{eff} for three algorithms viz. FPA, Tabular Q-Learning and DQL, all for the same episode. This episode generates the highest reward. Here we see that the FPA algorithm has the worst performance, which was expected since FPA has no power control or interference coordination. The tabular Q-Learning implementation has better performance compared with the FPA. This is because even though power control is introduced to the BSs, it not as effective, which explains why close to $\gamma_{\mathrm{eff}} = 9$ dB tabular Q-Learning underperforms FPA. Further, we observe that DQL, the proposed algorithm outperforms the tabular Q-Learning implementation, since DQL has resulted in a higher reward compared to tabular Q-Learning.This is because DQL has converged at a better

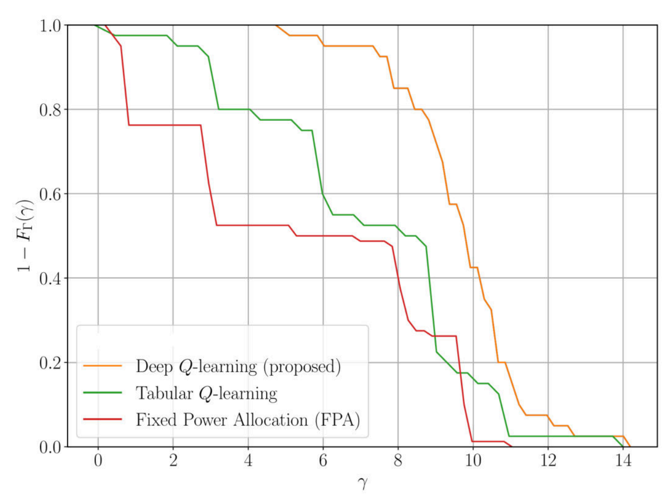

Fig. 5. CCDF plot of effective SINR for three different interference mitigation algorithms

solution, unlike the tabular Q-Learning, the convergence of which is hindered by the choice of a initialization of the state-action value function. However, as the effective SINR γ_{eff} approaches 13 dB, the UAVs are close to the BS center and therefore all power control algorithms perform almost similarly thereafter.

VI. CONCLUSION

With the unique challenges presented by the UAVs owing to their 3-D movement causing air-ground interference, new and unique methods to mitigate interference are required to address the challenges. In this paper, an algorithm is developed to maximize the downlink SINR in a 5G cellular network with UAVs. The UAVs experience interference from non-associated BSs due to multiple LoS channels. To this effect, power control, and interference mitigation algorithm using deep reinforcement learning is developed in this paper. By leveraging the cloud-based architecture of 5G systems, this algorithm is implemented in a cloud location and receives UAV measurements over the backhaul. The algorithm outperforms both the tabular Q -learning algorithm and the industry standard fixed power allocation algorithm. The proposed algorithm requires that the UAV sends its coordinates and its received SINR every millisecond to the base station. However, it does not require the knowledge of the channel state information, which removes the need for channel estimation and the associated training sequences. Moreover, the overall amount of feedback from the UAV is reduced because the UAV sends its coordinates and would not need to send explicit commands power control and interference mitigation. Thus, this algorithm provides a solution to one of the major roadblocks to implementing 5G-connected UAVs, i.e., air-ground interference

REFERENCES

[1] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, “A tutorial on UAVs for wireless networks: Applications, challenges, and

open problems," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[2] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, and E. Sousa, "5G communications race: Pursuit of more capacity triggers LTE in unlicensed band," *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 43–51, 2015.

[3] S. Al-Rubaye and A. Tsourdos, "Airport connectivity optimization for 5G ultra-dense networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 980–989, 2020.

[4] M. Ozpolat, S. Al-Rubaye, A. Williamson, and A. Tsourdos, "Integration of unmanned aerial vehicles and LTE: A scenario-dependent analysis," in *2022 International Conference on Connected Vehicle and Expo (ICCVE)*, 2022, pp. 1–6.

[5] A. S. Warrier, S. Al-Rubaye, D. Panagiotakopoulos, G. Inalhan, and A. Tsourdos, "Seamless handover in urban 5G-UAV systems using entropy weighted method," 2021.

[6] R. Amorim, P. Mogensen, T. Sorensen, I. Z. Kovacs, and J. Wigard, "Pathloss measurements and modeling for UAVs connected to cellular networks," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017, pp. 1–6.

[7] S. A. Busari, K. M. S. Huq, S. Mumtaz, J. Rodriguez, Y. Fang, D. C. Sicker, S. Al-Rubaye, and A. Tsourdos, "Generalized hybrid beamforming for vehicular connectivity using THz massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8372–8383, 2019.

[8] C. Li, S. C. Sun, S. Al-Rubaye, A. Tsourdous, and W. Guo, "Uncertainty propagation in neural network enabled multi-channel optimisation," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–4.

[9] S. Vasudevan, R. N. Pupala, and K. Sivanesan, "Dynamic eICIC — a proactive strategy for improving spectral efficiencies of heterogeneous LTE cellular networks by leveraging user mobility and traffic dynamics," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4956–4969, 2013.

[10] A. Chowdary, G. Chopra, A. Kumar, and L. R. Cenkeramaddi, "Enhanced user grouping and pairing scheme for coMP-NOMA-based cellular networks," in *2022 14th International Conference on COMmunication Systems NETworkS (COMSNETS)*, 2022, pp. 319–323.

[11] J. Pan, N. Ye, H. Yu, T. Hong, S. Al-Rubaye, S. Mumtaz, A. Al-Dulaimi, and I. Chih-Lin, "AI-driven blind signature classification for IoT connectivity: A deep learning approach," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.

[12] E. Chu, J. M. Kim, and B. C. Jung, "Interference analysis of directional UAV networks: A stochastic geometry approach," in *2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)*, 2019, pp. 9–12.

[13] L. Liu, S. Zhang, and R. Zhang, "Multi-beam UAV communication in cellular uplink: Cooperative interference cancellation and sum-rate maximization," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4679–4691, 2019.

[14] S. K. Gopalakrishnan, S. Al-Rubaye, and G. Inalhan, "Adaptive UAV swarm mission planning by temporal difference learning," in *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, 2021, pp. 1–10.

[15] C. Luo, J. Ji, Q. Wang, L. Yu, and P. Li, "Online power control for 5G wireless communications: A deep Q-network approach," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.

[16] H. S. Jang, H. Lee, and T. Q. S. Quek, "Deep learning-based power control for non-orthogonal random access," *IEEE Communications Letters*, vol. 23, no. 11, pp. 2004–2007, 2019.

[17] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1581–1592, 2020.

[18] H. Koumaras, G. Makropoulos, M. Batistatos, S. Kolometsos, A. Gogos, G. Xilouris, A. Sarlas, and M.-A. Kourtis, "5G-enabled UAVs with command and control software component at the edge for supporting energy efficient opportunistic networks," *Energies*, vol. 14, no. 5, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/5/1480

[19] E. T. S. Institution(ETSI), "System architecture for the 5G system-3GPP TS 23.501 v16 Rel16," 2020.

[20] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Foundations and Trends® in Machine Learning*, vol. 11, no. 3-4, pp. 219–354, 2018. [Online]. Available: http://dx.doi.org/10.1561/2200000071

2022-10-31

Interference mitigation for 5G-connected UAV using deep Q-learning framework

Warrier, Anirudh

IEEE

https://doi.org/10.1109/DASC55683.2022.9925817