

Multi-UAV IRS-assisted Communications: Multinode Channel Modeling and Fair Sum-Rate Optimization via Deep Reinforcement Learning

Giovanni Iacovelli ^{1,1}, Angelo Coluccia ², and Luigi Alfredo Grieco ²

¹Politecnico di Bari

²Affiliation not available

March 04, 2024

Abstract

Unmanned Aerial Vehicles (UAVs) combined with Intelligent Reflective Surfaces (IRSs) represent a cutting-edge technology for improving the channel capacity of wireless communications, by capitalizing on UAVs' 3D mobility coupled with the IRSs' smart radio capabilities. This work envisions a scenario in which a swarm of UAVs equipped with IRSs serves multiple Internet of Things (IoT) Ground Nodes (GNs) concurrently transmitting to a single Base Station (BS) via OFDMA. The huge number of passive elements composing the IRSs introduces a significant complexity in the mission design. Therefore, each IRS is divided into patches that can be simultaneously used to serve different nodes. Considering general Rician fading, a comprehensive channel model for IRS-assisted UAV-aided networks is derived. Then, a multi-objective mixed-integer non-linear programming problem is conceived to maximize the sum-rate of the GNs and, at the same time, minimize the difference among the users' data rates, by jointly optimizing the trajectories and the phase shift matrices. This non-convex problem, reformulated in terms of scheduling (i.e., patch-GN assignment), is challenging to solve. Hence, it is rearranged as a Markov Decision Process and a quasi-optimal solution is obtained via Deep Reinforcement Learning. Extensive simulation analysis is performed to validate the results and the accuracy of the proposed model.

Multi-UAV IRS-assisted Communications: Multinode Channel Modeling and Fair Sum-Rate Optimization via Deep Reinforcement Learning

Giovanni Iacovelli , *Member, IEEE*, Angelo Coluccia , *Senior Member, IEEE*
and Luigi Alfredo Grieco , *Senior Member, IEEE*

Abstract—Unmanned Aerial Vehicles (UAVs) combined with Intelligent Reflective Surfaces (IRSs) represent a cutting-edge technology for improving the channel capacity of wireless communications, by capitalizing on UAVs’ 3D mobility coupled with the IRSs’ smart radio capabilities. This work envisions a scenario in which a swarm of UAVs equipped with IRSs serves multiple Internet of Things (IoT) Ground Nodes (GNs) concurrently transmitting to a single Base Station (BS) via OFDMA. The huge number of passive elements composing the IRSs introduces a significant complexity in the mission design. Therefore, each IRS is divided into patches that can be simultaneously used to serve different nodes. Considering general Rician fading, a comprehensive channel model for IRS-assisted UAV-aided networks is derived. Then, a multi-objective mixed-integer non-linear programming problem is conceived to maximize the sum-rate of the GNs and, at the same time, minimize the difference among the users’ data rates, by jointly optimizing the trajectories and the phase shift matrices. This non-convex problem, reformulated in terms of scheduling (i.e., patch-GN assignment), is challenging to solve. Hence, it is rearranged as a Markov Decision Process and a quasi-optimal solution is obtained via Deep Reinforcement Learning. Extensive simulation analysis is performed to validate the results and the accuracy of the proposed model.

Index Terms—Internet of Drones, Intelligent Reflective Surface, Channel Modeling, Deep Reinforcement Learning.

I. INTRODUCTION

Internet of Drones (IoD) [1] is a disruptive paradigm which is expected to play a key role in upcoming 6G communications. The inherent versatility of Unmanned Aerial Vehicles (UAVs) enables a vast plethora of applications [2], such as sensing, delivery, surveying, and patrolling. Information

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, with particular reference to the partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”, CUP: D93C22000910001) and to the national center on “Sustainable Mobility” (CN00000023 - program “MOST”, CUP: D93C22000410001). It was also supported by the PRIN project no. 2017NS9FEY entitled “Realtime Control of 5G Wireless Networks: Taming the Complexity of Future Transmission and Computation Challenges” funded by the Italian MUR, by “The house of emerging technologies of Matera (CTEMT)” project funded by the Italian MIMIT, and by the PON AGREED projects (ARS01 00254) funded by the Italian MUR.

G. Iacovelli and L.A. Grieco are with the Department of Electrical and Information Engineering, Politecnico di Bari, Bari, Italy (email: giovanni.iacovelli@poliba.it, alfredo.grieco@poliba.it). A. Coluccia is with the Department of Engineering for Innovation, Università del Salento, Lecce, Italy (email: angelo.coluccia@unisalento.it). All are with the Consorzio Nazionale Interuniversitario per le Telecomunicazioni, Parma, Italy.

Copyright (c) 2023 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

exchange with ground infrastructures often plays a central role in these use cases. Moreover, UAVs can act as Flying Base Station (FBS) or relays, thus enhancing the communication capabilities of Internet of Things (IoT) networks [3], [4]. Indeed, IoT nodes are usually capability-constrained devices which can greatly benefit from the aid provided by drones’ versatility [5]–[7]. Another cornerstone of future cellular communications are Intelligent Reflective Surfaces (IRSs), also named Reconfigurable Intelligent Surfaces (RISs). These metasurfaces are composed by a matrix of elements, namely Passive Reflective Units (PRUs), and have demonstrated capabilities of controlling the radio environment. IRS-assisted wireless systems grant significant performance improvement, enhancing the channel quality perceived by communicating nodes. In particular, IRSs are able to shift incident electromagnetic waves by a programmable phase, thus yielding beamforming. This unique property is at the basis of the emerging concept of Smart Radio Environments [8].

Early literature has mainly considered to deploy IRSs on building facades, so as to reflect incoming signals towards multiple users. Recently, the possibility to equip UAVs with IRSs is attracting interest. Drones mobility adds more degrees of freedom that can be exploited to further improve the channel quality, with great advantage for IoT applications [9], [10]. Indeed, the high mobility of drones yields a better Line of Sight (LoS) link and a lower pathloss, due to the possibility to adjust the IRS location. Moreover, IRSs are characterized by a limited Size, Weight and Power consumption (SWaP) with respect to common phased-array antennas [11], which allows to prolong the mission duration, while still providing broad coverage. However, this comes with new challenges [12], [13] and in particular with the necessity of an accurate, general-purpose, and flexible channel model. Indeed, the presence of a Base Station (BS) with multiple IRSs and Ground Nodes (GNs), leads to interference patterns that need to be taken into account in case of system design and assessment.

This motivated recent studies that started to investigate the achievable performance of IRS-assisted UAV-aided systems. In [14] a non-convex optimization problem is formulated for the joint design of UAV trajectory, IRS’s phase shifts, scheduling, and resource allocation. The system employs Orthogonal Frequency Multiple Access (OFDMA), which introduces frequency and spatial selectivity in the fading of the resulting channel. However, as the majority of the present literature, it considers only the LoS component and hence a

deterministic channel which indeed is not able to capture the full characteristics of an actual wireless channel. In this regard, the literature on IRS has investigated different directions to obtain more realistic channel models. In [15] a formulation is proposed for IRS-aided wireless networks over Rician fading channels, which enables closed-form approximations of outage probability, average symbol error probability, and channel capacity. However, in this contribution the presence of a direct, yet weak, transmitter-receiver link is not considered. To fill this gap, [16] introduces a Rician model to account for the LoS, which yields a closed-form upper bound for the ergodic capacity, and a tight approximation for the outage probability. However, the resulting distribution cannot be recast as a known one, which limits the practical tractability. To partially circumvent this issue, simplified expressions are provided in [15], [16], but only for the asymptotic regime.

More tractable closed-form expressions considering Rician channels have been recently derived in [17], and evaluated in terms of outage and symbol error probability. However, the signal components are assumed to add coherently, which is not always strictly verified. In fact, there could be scenarios in which the same IRS has to serve multiple GN, thus introducing interference among surface elements. Besides, the presence of a direct, though weak, link is not considered, which would instead require to consider the additional interference between PRUs and GN. In order to account for the channel stochasticity, where interference between PRUs, BS and GNs is present, while considering the presence of a (possibly weak) GN-BS link, [18] proposes a channel gain approximation for IRS-assisted UAV-aided OFDMA communications. Although this contribution acknowledges several aspects of the communications, it does not explicitly model a multi-user multi-drone scenario. Moreover, the model introduces a huge number of degrees of freedom, one for each element of the IRS, which guarantees a great customization but at the same brings a tremendous complexity in terms of beam design. On the other hand, in [19] the number of degrees of freedom of the surface is drastically reduced by imposing a constraint on the phase shifts, which results to be extremely advantageous in terms of complexity but limits the surface to reflect the signal towards a single direction per time. Besides, [19] treats only a piece of the channel model, in a deterministic way, considering the sole path loss and not the stochastic term due to fading.

In light of the above, the major contributions given by this work are listed below.

- A comprehensive channel model for UAV-aided IRS-assisted OFDMA systems is derived. Specifically, a swarm of drones equipped with IRSs is considered. The latter are split into an arbitrary number of patches to simultaneously serve multiple GNs. The proposed formulation captures inherent dependencies on the relative positions of the actors (UAVs, GNs, BS) in the environment. To this aim, constructive/destructive interference patterns are explicitly modeled through a single controllable direction (two degrees of freedom) per patch. The derived expressions can be specialized in different cases as (i) absence of the direct BS-GN link, (ii) scattered (Rayleigh) BS-GN link, (iii) worst possible UAV-BS

alignment, and (iv) best possible UAV-BS alignment. Furthermore, the expression of the Signal-to-Noise Ratio (SNR) and maximum achievable data rate, given a certain outage probability, are derived and a theoretical analysis is provided.

- A Multi-Objective Mixed-Integer Non-Linear Programming (MINLP) problem is formulated starting from the proposed channel model. The aim is to maximize the sum-rate of GNs and, at the same time, minimize the differences among rates. Each node is served by a single BS (direct link) and by a swarm of drones (reflected links), thus requiring the jointly optimization of the phase shift matrices of the IRSs and the trajectories of the drones. To achieve this goal a suitable objective function is defined. It includes a fairness factor that determines the importance of a uniform distribution of the resources. The problem is then reformulated in terms of patch scheduling, i.e., patch-GN assignment. To overcome the intractability of the non-convex cosine patterns related to wave interference, the problem is finally rearranged as a Markov Decision Process (MDP) and solved via a Deep Reinforcement Learning (DRL) method i.e., a Proximal Policy Optimization (PPO)-based approach.
- An extensive simulation campaign is carried out to assess the validity of this work. First, the theoretical findings are corroborated by a thorough analysis which demonstrates, employing Monte Carlo simulations, the accuracy of the proposed channel model. Then, the solutions obtained by the DRL algorithm are analyzed for different number of patches and fairness levels. To further prove the effectiveness of the proposal, the obtained solution is compared with the special case of absence of the swarm and with a baseline approach.

The results of this work clearly indicate that (i) the presence of IRS-equipped drones strongly improves the channel capacity of the GNs, and (ii) the proposed solution outperforms the baseline in terms of total and average transmitted data.

The remainder of the present contribution is as follows: Section II describes the adopted system model. Section III presents the proposed channel model. Section IV describes the problem formulation and the proposed solution. Section V analyzes accuracy of the model and investigates the obtained numerical results. Finally, Section VI concludes the work and draws future research perspectives.

Notations adopted in this work are hereby described. Bold-face lower and capital case letters refer to vectors and matrices, respectively; $j = \sqrt{-1}$ is the imaginary unit; $\text{atan2}(x)$ denotes the four-quadrant arctangent of a real number x ; \mathbf{x}^\top is the transpose of a generic vector \mathbf{x} ; $\mathbf{x} \otimes \mathbf{y}$ denotes the Kronecker product between two generic vectors; $\text{diag}(\mathbf{x})$ represents a diagonal matrix whose diagonal is given by a vector \mathbf{x} ; $\arg(x)$ returns the phase of a complex number x ; $x \sim \mathcal{CN}(\mu, \sigma^2)$ define a circularly symmetric complex Gaussian distribution x with mean μ and variance σ^2 . For clarity, the adopted notations of this paper are summarized in Table I.

Symbol	Description	Symbol	Description
T	Total duration of the mission.	d_g^{BG}	Distance between the BS and g -th GN.
K	Number of discrete timesteps of the mission.	$d_{k,g,u,n,m}^{\text{RG}}$	Distance between the g -th GN and the (n,m) -th PRU.
U	Number of UAVs involved in the mission.	$d_{k,u,n,m}^{\text{BR}}$	Distance between the BS and (n,m) -th PRU.
G	Number of GNs deployed on the ground.	$\beta_{k,i,g}^{\text{BG}}$	The channel gain between the g -th GN and the BS.
I	Number of available subcarriers.	α^{BG}	The pathloss exponent for BS-GN link.
N	Number of PRUs on the rows.	r_g^{BG}	K-factor for BS-GN link.
M	Number of PRUs on the columns.	Ω_g^{BG}	Average power for BS-GN link.
δ_t	Duration of a timestep.	$\mathfrak{S}_{k,i,g,u,p_u}^{\text{RG}}$	The channel gain between a patch and a GN.
\mathbf{q}^{BS}	Location of the BS.	$\mathfrak{S}_{k,i,u,p_u}^{\text{BR}}$	The channel gain between a patch and the BS.
\mathbf{q}_g^{G}	Location of the g -th GN.	Φ_{k,u,p_u}	The IRS phase shift matrix of p_u -th patch of u -th UAV.
$\mathbf{q}_{k,u}^{\text{U}}$	Location of the u -th drone in the k -th timestep.	$\mathcal{G}_{k,i,g}$	The overall channel gain perceived by g -th GN.
$\mathbf{v}_{k,u}$	Speed of the u -th drone in the k -th timestep.	$C_{k,i,g}$	The channel capacity of the overall link of the g -th GN.
w	Area of a single PRU.	f_c	The carrier frequency.
$a_{k,u,m,n}$	The amplitude of the (n,m) -th element.	δ_f	The bandwidth of a single subcarrier.
$\phi_{k,u,m,n}$	The phase shift of the (n,m) -th element.	$\rho_{k,i,g}$	The transmit power of the g -th GN.
P_u	Number of patches of the u -th IRS.	$N_0\delta_f$	The thermal noise power.
E_{u,p_u}	Number of PRUs in p_u -th patch of the u -th IRS.		

TABLE I: Main notation adopted in this work.

II. SYSTEM MODEL

The envisioned scenario, depicted in Figure 1, considers a swarm of U IRS-equipped UAVs is in charge of optimally reflecting the incident signal coming from a set of G GNs to enhance the channel quality at a single BS, located at a known position $\mathbf{q}^{\text{BS}} = [x^{\text{B}} \ y^{\text{B}} \ z^{\text{B}}]^{\text{T}} \in \mathbb{R}^3$. Both GNs and the BS are equipped with a single-antenna communication apparatus. The entire mission lasts T seconds, split into K timeslots of δ_t seconds each. Therefore, the UAVs fly following trajectories discretized into K waypoints $\mathbf{q}_{k,u}^{\text{U}} = [x_{k,u}^{\text{U}} \ y_{k,u}^{\text{U}} \ z_{k,u}^{\text{U}}]^{\text{T}} \in \mathbb{R}^3$ with a speed denoted by $\mathbf{v}_{k,u} = [v_{k,u}^x \ v_{k,u}^y \ v_{k,u}^z]^{\text{T}} \in \mathbb{R}^3$, $u = 1, \dots, U$, $k = 1, \dots, K$. IRSs are composed by $N \times M$ PRUs, all with the same size $w = d^x \times d^y \text{ m}^2$. The (n,m) -th element (i) has a center defined as $[(m - \frac{1}{2})d^x, (n - \frac{1}{2})d^y]^{\text{T}}$ with $m = 1 - \frac{M}{2}, \dots, \frac{M}{2}, n = 1 - \frac{N}{2}, \dots, \frac{N}{2}$ and (ii) reflects the incident signal through a complex factor $\Gamma_{k,u,n,m} = a_{k,u,n,m} e^{j\phi_{k,u,n,m}}$, where $a_{k,u,n,m}$ is the amplitude and $\phi_{k,u,n,m} \in [-\pi, \pi)$ is the phase shift. The IRSs are assumed to be controlled by the BS through a dedicated control channel. PRUs of the u -th IRS are grouped into P_u patches of $E_{u,p_u} = E_{u,p_u}^{\text{R}} \times E_{u,p_u}^{\text{C}} = (R_{u,p_u} - r_{u,p_u}) \times (C_{u,p_u} - c_{u,p_u})$ elements, $p_u = 1, \dots, P_u$, where $1 - \frac{M}{2} \leq r_{u,p_u} \leq R_{u,p_u} \leq \dots \leq r_{u,p_u} \leq R_{u,p_u} \leq \frac{M}{2}$ and $1 - \frac{N}{2} \leq c_{u,p_u} \leq C_{u,p_u} \leq \dots \leq c_{u,p_u} \leq C_{u,p_u} \leq \frac{N}{2}$. Besides, GNs are located at known positions $\mathbf{q}_g^{\text{G}} = [x_g^{\text{G}} \ y_g^{\text{G}} \ z_g^{\text{G}}]^{\text{T}} \in \mathbb{R}^3$, $g = 1, \dots, G$, leading to the definition of the distance $d_g^{\text{BG}} = \|\mathbf{q}^{\text{BS}} - \mathbf{q}_g^{\text{G}}\|$ from the BS, which does not change over time*. On the opposite, UAV-GN and BS-UAV distances depend on k -th timeslot since they are a function of the time-varying UAV positions. In particular, in far-field, the distance between each surface element of the u -th drone and the g -th GN can be approximated as [19]

$$d_{k,g,u,n,m}^{\text{RG}} \simeq d_{k,g,u}^{\text{RG}} - \left(m - \frac{1}{2}\right) d^x \sin \theta_{k,g,u}^{\text{RG}} \cos \varphi_{k,g,u}^{\text{RG}} + \left(n - \frac{1}{2}\right) d^y \sin \theta_{k,g,u}^{\text{RG}} \sin \varphi_{k,g,u}^{\text{RG}}, \quad (1)$$

where $d_{k,g,u}^{\text{RG}}$ is the distance from the center of the IRS, while $\theta_{k,g,u}^{\text{RG}}$ and $\varphi_{k,g,u}^{\text{RG}}$ denote the vertical and horizontal

*In this work, since the kinetics-related parameters are known to the BS, the Doppler effect is assumed to be perfectly compensated.

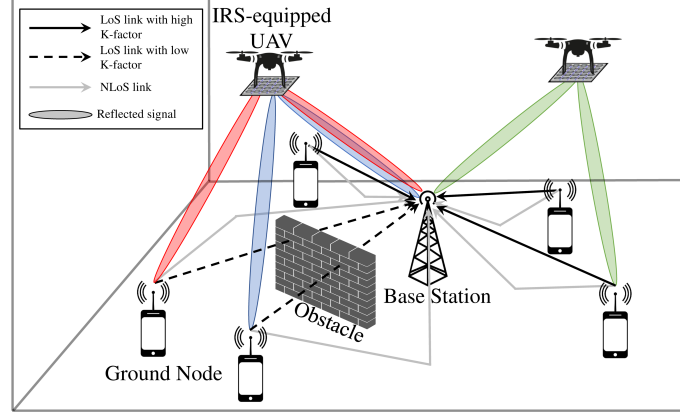


Fig. 1: Reference scenario.

Angles of Arrival (AoA) of the signal. In particular, $\theta_{k,g,u}^{\text{RG}} = \text{acos} \left(\frac{z_g^{\text{G}} - z_{k,u}^{\text{U}}}{d_{k,g,u}^{\text{RG}}} \right)$ and $\varphi_{k,g,u}^{\text{RG}} = \text{atan2} \left(\frac{y_g^{\text{G}} - y_{k,u}^{\text{U}}}{x_g^{\text{G}} - x_{k,u}^{\text{U}}} \right)$. Similarly, the distance between the BS and each PRU of the u -th drone is

$$d_{k,u,n,m}^{\text{BR}} \simeq d_{k,u}^{\text{BR}} - \left(m - \frac{1}{2}\right) d^x \sin \theta_{k,u}^{\text{BR}} \cos \varphi_{k,u}^{\text{BR}} + \left(n - \frac{1}{2}\right) d^y \sin \theta_{k,u}^{\text{BR}} \sin \varphi_{k,u}^{\text{BR}}, \quad (2)$$

where $d_{k,u}^{\text{BR}}$ is defined as the distance from the center of the IRS, while $\theta_{k,u}^{\text{BR}}$ and $\varphi_{k,u}^{\text{BR}}$ are the vertical and horizontal Angles of Departure (AoD), so that $\theta_{k,g,u}^{\text{BR}} = \text{acos} \left(\frac{z_g^{\text{G}} - z_{k,u}^{\text{U}}}{d_{k,g,u}^{\text{BR}}} \right)$ and $\varphi_{k,g,u}^{\text{BR}} = \text{atan2} \left(\frac{y_g^{\text{G}} - y_{k,u}^{\text{U}}}{x_g^{\text{G}} - x_{k,u}^{\text{U}}} \right)$.

III. PROPOSED CHANNEL MODELING

The whole communication system employs OFDMA which allows to mitigate the interference among the network entities. As a consequence, the whole available bandwidth B is split into I subcarriers of δ_f Hz each that can be assigned to GNs to transmit. Moreover, the power radiation pattern functions, including antenna gains, which characterize the antennas of GNs, BS, and IRSs are denoted by $F^{\text{GN}}(\theta, \varphi), \forall g, F^{\text{BS}}(\theta, \varphi),$

$$\mathbf{h}_{k,i,g,u,p_u}^{\text{RG, LoS}} = \left[e^{j2\pi f_i \frac{(c_{u,p_u} - \frac{1}{2})d^X \sin \theta_{k,g,u}^{\text{RG}} \cos \varphi_{k,g,u}^{\text{RG}}}{c}} \dots e^{j2\pi f_i \frac{(C_{u,p_u} - \frac{1}{2})d^X \sin \theta_{k,g,u}^{\text{RG}} \cos \varphi_{k,g,u}^{\text{RG}}}{c}} \right]^T \otimes \left[e^{j2\pi f_i \frac{(r_{u,p_u} - \frac{1}{2})d^Y \sin \theta_{k,g,u}^{\text{RG}} \sin \varphi_{k,g,u}^{\text{RG}}}{c}} \dots e^{j2\pi f_i \frac{(R_{u,p_u} - \frac{1}{2})d^Y \sin \theta_{k,g,u}^{\text{RG}} \sin \varphi_{k,g,u}^{\text{RG}}}{c}} \right]^T, \quad (8)$$

$$\mathbf{h}_{k,i,u,p_u}^{\text{BR, LoS}} = \left[e^{j2\pi f_i \frac{(c_{u,p_u} - \frac{1}{2})d^X \sin \theta_{k,u}^{\text{BR}} \cos \varphi_{k,u}^{\text{BR}}}{c}} \dots e^{j2\pi f_i \frac{(C_{u,p_u} - \frac{1}{2})d^X \sin \theta_{k,u}^{\text{BR}} \cos \varphi_{k,u}^{\text{BR}}}{c}} \right]^T \otimes \left[e^{j2\pi f_i \frac{(r_{u,p_u} - \frac{1}{2})d^Y \sin \theta_{k,u}^{\text{BR}} \sin \varphi_{k,u}^{\text{BR}}}{c}} \dots e^{j2\pi f_i \frac{(R_{u,p_u} - \frac{1}{2})d^Y \sin \theta_{k,u}^{\text{BR}} \sin \varphi_{k,u}^{\text{BR}}}{c}} \right]^T, \quad (12)$$

$$G_{k,i,g,u,p_u} = \mathbf{g}_{k,i,u,p_u}^{\text{BR}} \mathbf{\Phi}_{k,u,p_u} \mathbf{g}_{k,i,g,u,p_u}^{\text{RG}} = \eta_{k,g,u} \sum_{n=r_{u,p_u}}^{R_{u,p_u}} \sum_{m=c_{u,p_u}}^{C_{u,p_u}} \underbrace{h_{k,i,u,p_u,n,m}^{\text{BR}} \mathbf{\Gamma}_{k,u,p_u,n,m} h_{k,i,g,u,p_u,n,m}^{\text{RG}}}_{h_{k,i,g,u,p_u,n,m}^{\text{BRG}}} \quad (15)$$

and $F^{\text{IRS}}(\theta, \varphi)$, respectively. As known, these functions define how the transmitted/received power at each antenna varies along a certain direction in inclination θ and azimuth φ angles.

The channel gain between the g -th GN and the BS, in k -th timeslot and subcarrier $i = 1, \dots, I$ is

$$g_{k,i,g}^{\text{BG}} = \sqrt{\beta^{\text{BG}} d_g^{\text{BG} - \alpha^{\text{BG}}} F^{\text{BG}}(\theta_g^{\text{BG}}, \varphi_g^{\text{BG}}) h_{k,i,g}^{\text{BG}}}, \quad (3)$$

where β^{BG} denotes the channel power gain at the reference distance of 1 m, α^{BG} is the pathloss exponent, $F^{\text{BG}}(\cdot, \cdot) = F^{\text{BS}}(\cdot, \cdot) F^{\text{GN}}(\cdot, \cdot)$, and $h_{k,i,g}^{\text{BG}} \sim \mathcal{CN}(\mu_{k,i,g}^{\text{BG}}, 2\sigma_{k,i,g}^{\text{BG}^2})$ denotes the channel coefficient representing the stochastic fluctuation due to multi-path propagation and fading. It is worth noting that the channel envelope $|h_{k,i,g}^{\text{BG}}|$ is generally Rician [20], with K-factor $\kappa_g^{\text{BG}} = \frac{|\mu_{k,i,g}^{\text{BG}}|^2}{2\sigma_{k,i,g}^{\text{BG}^2}}$ and average power $\Omega_g^{\text{BG}} = |\mu_{k,i,g}^{\text{BG}}|^2 + 2\sigma_{k,i,g}^{\text{BG}^2} = 1$, since the presence of a (possibly weak) LoS path is taken into account — according to the propagation characteristics of the actual environment and the presence of obstacles, as depicted in Figure 1. Therefore, the channel coefficient can be modeled as

$$h_{k,i,g}^{\text{BG}} = \sqrt{\frac{\kappa_g^{\text{BG}}}{\kappa_g^{\text{BG}} + 1}} \bar{h}_{k,i,g}^{\text{BG}} + \sqrt{\frac{1}{\kappa_g^{\text{BG}} + 1}} \tilde{h}_{k,i,g}^{\text{BG}}. \quad (4)$$

In particular, $\bar{h}_{k,i,g}^{\text{BG}} = e^{-j2\pi f_i \frac{d_{k,i,g}^{\text{BG}}}{c}}$ is the LoS component, characterized by a phase shift depending on the user's subcarrier index, $f_i = f_c + i\delta_f$, f_c is the carrier frequency, and $\tilde{h}_{k,i,g}^{\text{BG}} \sim \mathcal{CN}(0, 1)$ describes the small-scale fading resulting from NLoS propagation.

The channel gain between the p_u -th patch of the u -th UAV and the g -th GN in timeslot k , subcarrier i , is obtained by collecting in a vector the corresponding E_{u,p_u} channel gains of the PRUs, i.e.

$$\mathbf{g}_{k,i,g,u,p_u}^{\text{RG}} = \sqrt{\beta^{\text{RG}} d_{k,g,u}^{\text{RG} - \alpha^{\text{RG}}} F^{\text{RG}}(\theta_{k,g,u}^{\text{RG}}, \varphi_{k,g,u}^{\text{RG}}) \mathbf{h}_{k,i,g,u,p_u}^{\text{RG}}}, \quad (5)$$

$$\mathbf{h}_{k,i,g,u,p_u}^{\text{RG}} = \sqrt{\frac{\kappa_{k,g,u}^{\text{RG}}}{\kappa_{k,g,u}^{\text{RG}} + 1}} \bar{\mathbf{h}}_{k,i,g,u,p_u}^{\text{RG}} + \sqrt{\frac{1}{\kappa_{k,g,u}^{\text{RG}} + 1}} \tilde{\mathbf{h}}_{k,i,g,u,p_u}^{\text{RG}}, \quad (6)$$

$$\bar{\mathbf{h}}_{k,i,g,u,p_u}^{\text{RG}} = e^{-j2\pi f_i \frac{d_{k,i,g,u}^{\text{RG}}}{c}} \mathbf{h}_{k,i,g,u,p_u}^{\text{RG, LoS}}, \quad (7)$$

$\tilde{\mathbf{h}}_{k,i,g,u,p_u}^{\text{RG}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{E_{u,p_u}})$, and $F^{\text{RG}}(\cdot, \cdot) = F^{\text{IRS}}(\cdot, \cdot) F^{\text{GN}}(\cdot, \cdot)$. Moreover, $\mathbf{h}_{k,i,g,u,p_u}^{\text{RG, LoS}} \in \mathbb{C}^{E_{u,p_u} \times 1}$ in (7) denotes the far-field array response defined in (8) (reported at the top of this page).

Similarly, the channel gain between a certain patch p_u of a UAV u and the BS is

$$\mathbf{g}_{k,i,u,p_u}^{\text{BR}} = \sqrt{\beta^{\text{BR}} d_{k,u}^{\text{BR} - \alpha^{\text{BR}}} F^{\text{BR}}(\theta_{k,u}^{\text{BR}}, \varphi_{k,u}^{\text{BR}}) \mathbf{h}_{k,i,u,p_u}^{\text{BR}}}, \quad (9)$$

$$\mathbf{h}_{k,i,u,p_u}^{\text{BR}} = \sqrt{\frac{\kappa_{k,u}^{\text{BR}}}{\kappa_{k,u}^{\text{BR}} + 1}} \bar{\mathbf{h}}_{k,i,u,p_u}^{\text{BR}} + \sqrt{\frac{1}{\kappa_{k,u}^{\text{BR}} + 1}} \tilde{\mathbf{h}}_{k,i,u,p_u}^{\text{BR}}, \quad (10)$$

$$\bar{\mathbf{h}}_{k,i,u,p_u}^{\text{BR}} = e^{-j2\pi f_i \frac{d_{k,i,u}^{\text{BR}}}{c}} \mathbf{h}_{k,i,u,p_u}^{\text{BR, LoS}}, \quad (11)$$

where $\mathbf{h}_{k,i,u,p_u}^{\text{BR, LoS}} \in \mathbb{C}^{E_{u,p_u} \times 1}$ is given in (12) (reported at the top of this page), $F^{\text{BR}}(\cdot, \cdot) = F^{\text{BS}}(\cdot, \cdot) F^{\text{IRS}}(\cdot, \cdot)$, and $\tilde{\mathbf{h}}_{k,i,u,p_u}^{\text{BR}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{E_{u,p_u}})$. Note that $\kappa_{k,g,u}^{\text{BR}}$ depends on the UAV, specifically on the elevation angle $\theta_{k,g,u}^{\text{BR}} = \frac{\pi}{2} - \theta_{k,g,u}^{\text{BR}}$ [21]:

$$\kappa_{k,u}^{\text{BR}} = \kappa_{k,u}^{\text{MIN, BR}} e^{\gamma^{\text{BR}} \theta_{k,g,u}^{\text{BR}}}, \quad \gamma^{\text{BR}} = \frac{2}{\pi} \ln \frac{\kappa_{k,u}^{\text{MAX, BR}}}{\kappa_{k,u}^{\text{MIN, BR}}}, \quad (13)$$

with $\kappa_{k,g,u}^{\text{MIN, BR}}$ and $\kappa_{k,g,u}^{\text{MAX, BR}}$ the minimum and maximum possible K-factors, respectively. The same holds true for $\kappa_{k,g,u}^{\text{RG}}$ and $\kappa_{k,g,u}^{\text{BG}}$. The IRS phase shift matrix $\mathbf{\Phi}_{k,u,p_u} \in \mathbb{C}^{E_{u,p_u} \times E_{u,p_u}}$, for each timeslot k , UAV u , patch p_u is defined as

$$\begin{aligned} \mu_{G_{k,i,g,u,p_u}} &= a_{k,u,p_u} \eta_{k,g,u} \sqrt{\kappa_{k,g,u}^{\text{BRG}}} e^{-j \left(\frac{2\pi f_i}{c} (d_{k,u}^{\text{BR}} + d_{k,g,u}^{\text{RG}}) \right)} \sum_{m=C_{u,p_u}}^{C_{u,p_u}} e^{j \frac{2\pi f_i}{c} d^X (m - \frac{1}{2})} \psi_{k,g,u,p_u}^X \sum_{n=R_{u,p_u}}^{R_{u,p_u}} e^{j \frac{2\pi f_i}{c} d^Y (n - \frac{1}{2})} \psi_{k,g,u,p_u}^Y \\ &= a_{k,u,p_u} \eta_{k,g,u} \sqrt{\kappa_{k,g,u}^{\text{BRG}}} \underbrace{\frac{\sin \left(\frac{\pi f_i}{c} d^X E_{u,p_u}^C \psi_{k,g,u,p_u}^X \right) \sin \left(\frac{\pi f_i}{c} d^Y E_{u,p_u}^R \psi_{k,g,u,p_u}^Y \right)}{\sin \left(\frac{\pi f_i}{c} d^X \psi_{k,g,u,p_u}^X \right) \sin \left(\frac{\pi f_i}{c} d^Y \psi_{k,g,u,p_u}^Y \right)}}_{\chi_{k,g,u,p_u}} \underbrace{e^{-j \left(\frac{2\pi f_i}{c} (d_{k,u}^{\text{BR}} + d_{k,g,u}^{\text{RG}}) \right)}}_{\omega_{k,i,g,u}} \end{aligned} \quad (19)$$

$$\begin{aligned} \nu_{k,i,g}^2 &= \sum_{u=1}^U \sum_{p_u=1}^{P_u} |\mu_{G_{k,i,g,u,p_u}}|^2 + |\mu_{i,g}^{\text{BG}}|^2 + 2 \sum_{u \geq u'} \sum_{p_u > p'_u} \Re(\mu_{G_{k,i,g,u,p_u}}^* \mu_{G_{k,i,g,u',p'_u}}) + 2 \sum_{u=1}^U \sum_{p_u=1}^{P_u} \Re(\mu_{G_{k,i,g,u,p_u}}^* \mu_{i,g}^{\text{BG}}) \\ &= \underbrace{\sum_{u=1}^U \sum_{p_u=1}^{P_u} \chi_{k,i,g,u,p_u}^2 + 2 \sum_{u \geq u'} \sum_{p_u > p'_u} |\chi_{k,i,g,u,p_u}| |\chi_{k,i,g,u',p'_u}| \cos(\varphi_{k,i,g,u} - \varphi_{k,i,g,u'})}_{\nu_{k,i,g}^{\text{BRG}^2}} \\ &\quad + \underbrace{\lambda_g^2 \kappa_g^{\text{BG}} + 2 \sum_{u=1}^U \sum_{p_u=1}^{P_u} |\chi_{k,i,g,u,p_u}| |\lambda_g \sqrt{\kappa_g^{\text{BG}}}| \cos\left(\varphi_{k,i,g,u} + 2\pi f_i \frac{d_g^{\text{BG}}}{c}\right)}_{\nu_{k,i,g}^{\text{BG}^2}} \end{aligned} \quad (23)$$

$$\Phi_{k,u,p_u} = \text{diag} \left(\Gamma_{k,u,r_{u,p_u},C_{u,p_u}}, \dots, \Gamma_{k,u,r_{u,p_u},C_{u,p_u}}, \dots, \Gamma_{k,u,R_{u,p_u},C_{u,p_u}}, \dots, \Gamma_{k,u,R_{u,p_u},C_{u,p_u}} \right) \quad (14)$$

recalling that in general $\Gamma_{k,u,n,m} = a_{k,u,n,m} e^{j\phi_{k,u,n,m}}$. Since each patch p_u is meant to coherently reflect the incident signal towards a certain location (to serve one of the G GNs), all its elements can share the same amplitude $a_{k,u,n,m} = a_{k,u,p_u}$ for all (n, m) belonging to patch p_u , and their phases can be described in terms of two parameters only, denoted by ϕ_{k,u,p_u}^X and ϕ_{k,u,p_u}^Y . In general, it is possible to reduce the amount of degrees of freedom introduced by PRUs by imposing that:

$$\frac{2\pi f_i}{c} \left(d^X \left(m - \frac{1}{2} \right) \phi_{k,u,p_u}^X + d^Y \left(n - \frac{1}{2} \right) \phi_{k,u,p_u}^Y \right) = \phi_{k,u,n,m}.$$

By considering (5), (9), and (14), the composite channel gain G_{k,i,g,u,p_u} is given by (15) (reported at the top of the previous page) where $h_{k,i,g,u,p_u,n,m}^{\text{BR}}$ and $h_{k,i,g,u,p_u,n,m}^{\text{RG}}$ are the components of the channel vectors $\mathbf{h}_{k,i,g,u,p_u}^{\text{BR}}$ and $\mathbf{h}_{k,i,g,u,p_u}^{\text{RG}}$, respectively. Moreover, $\eta_{k,g,u} = \sqrt{\beta^{\text{BRG}} d_{k,u}^{\text{BR}} - \alpha^{\text{BR}} d_{k,g,u}^{\text{RG}} - \alpha^{\text{RG}} F^{\text{BRG}}(\theta_{k,u}^{\text{BR}}, \varphi_{k,u}^{\text{BR}}, \theta_{k,g,u}^{\text{RG}}, \varphi_{k,g,u}^{\text{RG}})}$, $\lambda_g = \sqrt{\beta^{\text{BG}} d_g^{\text{BG}} - \alpha^{\text{BG}} F^{\text{BG}}(\theta_g^{\text{BG}}, \varphi_g^{\text{BG}})}$, $F^{\text{BRG}}(\cdot, \cdot, \cdot, \cdot) = F^{\text{BR}}(\cdot, \cdot) F^{\text{RG}}(\cdot, \cdot)$, $\beta^{\text{BRG}} = \beta^{\text{BR}} \beta^{\text{RG}}$, and $h_{k,i,g,u,p_u,n,m}^{\text{BRG}}$ describes the composite BS-PRU-GN channel coefficient.

The composite BS-PRU-GN channel coefficient involves the pairwise product of two complex Gaussian Random Variables (RVs), whose distribution, named complex double Gaussian, is given in terms of an infinite sum of modified Bessel functions [22]. For better tractability, it is however possible to approximate such a product through a complex Gaussian [18], thus implying that the envelope $|h_{k,i,g,u,p_u,n,m}^{\text{BRG}}|$ is a Rician

RV. The following expression for the channel coefficient, for the generic (n, m) -the element of the p_u -th patch, is obtained:

$$\begin{aligned} h_{k,i,g,u,p_u,n,m}^{\text{BRG}} &= a_{k,u,p_u} \left(\sqrt{\kappa_{k,g,u}^{\text{BRG}}} e^{-j \frac{2\pi f_i}{c} \Psi_{k,i,u,p_u,n,m}} \right. \\ &\quad \left. + \sqrt{\tilde{\kappa}_{k,g,u}^{\text{BRG}}} \tilde{h}^{\text{BRG}} \right), \end{aligned} \quad (16)$$

where

$$\kappa_{k,g,u}^{\text{BRG}} = \frac{\kappa_{k,u}^{\text{BR}} \kappa_{k,g,u}^{\text{RG}}}{(\kappa_{k,u}^{\text{BR}} + 1)(\kappa_{k,g,u}^{\text{RG}} + 1)}, \quad (17)$$

$$\tilde{\kappa}_{k,g,u}^{\text{BRG}} = \frac{\kappa_{k,u}^{\text{BR}} + \kappa_{k,g,u}^{\text{RG}}}{(\kappa_{k,u}^{\text{BR}} + 1)(\kappa_{k,g,u}^{\text{RG}} + 1)}, \quad (18)$$

$$\begin{aligned} \Psi_{k,i,g,u,p_u,n,m} &= d_{k,u}^{\text{BR}} + d_{k,g,u}^{\text{RG}} - d^X \left(m - \frac{1}{2} \right) \psi_{k,g,u,p_u}^X \\ &\quad - d^Y \left(n - \frac{1}{2} \right) \psi_{k,g,u,p_u}^Y, \end{aligned}$$

$$\begin{aligned} \psi_{k,g,u,p_u}^X &= \sin \theta_{k,u}^{\text{BR}} \cos \varphi_{k,u}^{\text{BR}} + \sin \theta_{k,g,u}^{\text{RG}} \cos \varphi_{k,g,u}^{\text{RG}} + \phi_{k,u,p_u}^X, \\ \psi_{k,g,u,p_u}^Y &= \sin \theta_{k,u}^{\text{BR}} \sin \varphi_{k,u}^{\text{BR}} + \sin \theta_{k,g,u}^{\text{RG}} \sin \varphi_{k,g,u}^{\text{RG}} + \phi_{k,u,p_u}^Y \end{aligned}$$

and $\tilde{h}^{\text{BRG}} \sim \mathcal{CN}(0, 1)$ since phase terms are irrelevant to the scatter component [14].

The expression derived in (15), which includes (16) and the subsequent definitions, requires further elaboration in order to obtain a more compact form, with a lower computational complexity. To this aim, the following result will be useful to compute the sum of the PRUs' channel coefficients belonging to the same patch p_u .

Lemma 1. *The finite sum of complex exponentials $e^{j(l-\frac{1}{2})x}$ for $l = 1 - \frac{L}{2}, \dots, \frac{L}{2}$, with L even, can be rearranged as*

a ratio of sine (or equivalently as L times a ratio of sinc) functions, as

$$\sum_{l=1-\frac{l}{2}}^{\frac{l}{2}} e^{j(l-\frac{1}{2})x} = \frac{\sin(\frac{L}{2}x)}{\sin(\frac{x}{2})} = L \frac{\text{sinc}(\frac{L}{2}x)}{\text{sinc}(\frac{x}{2})}$$

Proof. Please refer to [19] and [23]. ■

The following theorem is one of the main contributions of this work. Indeed, it characterizes the overall gain $G_{k,i,g}$, given by the sum of all signals reflected towards user g (by any patch from any UAV) plus the direct link.

Theorem 1. Let $G_{k,i,g}$ be the channel gain perceived by a GN g , in timeslot k and subcarrier i . The channel envelope $|G_{k,i,g}|$ is a Rician RV having $\kappa_{k,i,g} = \frac{\nu_{k,i,g}^2}{2\sigma_{G_{k,i,g}}^2}$ and average power $\Omega_{k,i,g} = \nu_{k,i,g}^2 + 2\sigma_{G_{k,i,g}}^2$, with $\nu_{k,i,g}^2$ and $2\sigma_{G_{k,i,g}}^2$ defined in (23) and (24), respectively.

Proof. Leveraging the result of Lemma 1, since (15) is a sum of complex Gaussian distributions, and assuming that all the elements of the same patch p_u have the same amplitude a_{k,u,p_u} , the channel coefficients sum of all the PRUs $G_{k,i,g,u,p_u} \sim \mathcal{CN}(\mu_{G_{k,i,g,u,p_u}}, 2\sigma_{G_{k,i,g,u,p_u}}^2)$ is described by the mean in (19) (reported at the top of the previous page) and by the second moment

$$2\sigma_{G_{k,i,g,u,p_u}}^2 = a_{k,u,p_u}^2 E_{u,p_u} \eta_{k,g,u}^2 \tilde{\kappa}_{k,g,u}^{\text{BRG}}. \quad (20)$$

Finally, the expression of the channel gain perceived by a certain GN g is

$$G_{k,i,g} = \sum_{u=1}^U \sum_{p_u=1}^{P_u} G_{k,i,g,u,p_u} + g_{k,i,g}^{\text{BG}}. \quad (21)$$

Following the same rationale adopted in [18] it is possible to derive the K-factor $\kappa_{k,i,g}$ and average power $\Omega_{k,i,g}$ of the envelope of $G_{k,i,g}$ as

$$\kappa_{k,i,g} = \frac{\nu_{k,i,g}^2}{2\sigma_{G_{k,i,g}}^2}, \quad \Omega_{k,i,g} = \nu_{k,i,g}^2 + 2\sigma_{G_{k,i,g}}^2, \quad (22)$$

being $\nu_{k,i,g}^2 \triangleq |\mu_{G_{k,i,g}}|^2$ the squared LoS component, defined in (23) (reported at the top of the previous page) with $\varphi_{k,i,g,u} = \arg(\omega_{k,i,g,u})$, $\forall k, i, g, u$ and $\bar{\kappa}_g^{\text{BG}} = \frac{\kappa_g^{\text{BG}}}{\kappa_g^{\text{BG}} + 1}$. Finally, the Non Line of Sight (NLoS) component reads

$$2\sigma_{G_{k,i,g}}^2 = \underbrace{\sum_{u=1}^U \eta_{k,g,u}^2 \tilde{\kappa}_{k,g,u}^{\text{BRG}} \sum_{p_u=1}^{P_u} a_{k,u,p_u}^2 E_{u,p_u}}_{2\sigma_{G_{k,i,g}}^{\text{BRG}}{}^2} + \underbrace{\lambda_g^2 \bar{\kappa}_g^{\text{BG}}}_{2\sigma_{G_{k,i,g}}^{\text{BG}}{}^2}, \quad (24)$$

where $\tilde{\kappa}_g^{\text{BG}} = (\kappa_g^{\text{BG}} + 1)^{-1}$. ■

Corollary 1. The proposed general channel model can be specialized in case of scattered, i.e., Rayleigh fading ($\kappa_g^{\text{BG}} \rightarrow 0$), or even totally absent direct BS-GN link ($\lambda_g \rightarrow 0$). In both cases, the squared LoS component $\nu_{k,i,g}^2 \rightarrow \nu_{k,i,g}^{\text{BRG}}{}^2$, since $\nu_{k,i,g}^{\text{BG}}{}^2 \rightarrow 0$. However, the NLoS component is $2\sigma_{G_{k,i,g}}^2 \rightarrow 2\sigma_{G_{k,i,g}}^{\text{BRG}}{}^2 + \lambda_g^2$

in the former case, while is $2\sigma_{G_{k,i,g}}^2 \rightarrow 2\sigma_{G_{k,i,g}}^{\text{BRG}}{}^2$ in the latter case since $2\sigma_{G_{k,i,g}}^{\text{BRG}}{}^2 \rightarrow 0$.

Based on the above results, it is possible to obtain expressions for the maximum achievable data rates, outage probability, and other inherent performance metrics. In particular, recalling Shannon's capacity formula, the channel capacity in timeslot k and subcarrier i is given by

$$C_{k,i,g} = \delta_F \log_2 \left(1 + \frac{\rho_{k,i,g} |G_{k,i,g}|^2}{N_0 \delta_F} \right), \quad (25)$$

where $\rho_{k,i,g}$ is the transmit power of GNs in i -th subcarrier and N_0 is the spectral noise power. Given a maximum achievable data rate $R_{k,i,g}$, to guarantee a reliable communication, it is required that the outage probability remains below a threshold ε , such that

$$\begin{aligned} \mathbb{P}(C_{k,i,g} < R_{k,i,g}) &= \mathbb{P} \left(|G_{k,i,g}|^2 < \frac{N_0 \delta_F (2^{\frac{R_{k,i,g}}{\delta_F}} - 1)}{\rho_{k,i,g}} \right) \\ &= F_{|G_{k,i,g}|^2} \left(\underbrace{\frac{N_0 \delta_F (2^{\frac{R_{k,i,g}}{\delta_F}} - 1)}{\rho_{k,i,g}}}_u \right) \leq \varepsilon, \quad \forall \begin{matrix} k = 1, \dots, K, \\ i = 1, \dots, I, \end{matrix} \end{aligned}$$

with $F_{|G_{k,i,g}|^2}$ denoting the Cumulative Distribution Function (CDF) of $|G_{k,i,g}|^2$. Therefore, considering the maximum tolerable outage, i.e., ε , the previous equation can be rewritten as

$$F_{|G_{k,i,g}|^2}(u) = 1 - Q_m \left(\sqrt{2\kappa_{k,i,g}}, \sqrt{\frac{2(\kappa_{k,i,g} + 1)}{\Omega_{k,i,g}}} u \right) = \varepsilon,$$

with Q_m denoting the Marcum Q-function. It follows that the inverse expression, with respect to the second argument is:

$$\sqrt{\frac{2(\kappa_{k,i,g} + 1)}{\Omega_{k,i,g}}} u = Q_m^{-1}(\sqrt{2\kappa_{k,i,g}}, 1 - \varepsilon) \triangleq \zeta_{k,i,g}(\varepsilon).$$

Since the inverse Marcum Q-function has no closed-form expression, to avoid numerical computation an approximated formula can be used as in [24, Eq. 17]

$$\zeta_{k,i,g}(\varepsilon) = \begin{cases} \sqrt{-2 \log(1 - \varepsilon)} e^{\frac{\kappa_{k,i,g}}{2}}, & \text{for } \kappa_{k,i,g} \leq \frac{K_0^2}{2} \\ \sqrt{2\kappa_{k,i,g} + \frac{1}{2Q^{-1}(\varepsilon)}} \\ \times \log\left(\frac{\sqrt{2\kappa_{k,i,g}}}{\sqrt{2\kappa_{k,i,g}} - Q^{-1}(\varepsilon)}\right) - Q^{-1}(\varepsilon), & \text{for } \kappa_{k,i,g} \geq \frac{K_0^2}{2} \end{cases}$$

with K_0 the intersection of the sub-functions at $\sqrt{2\kappa_{k,i,g}} > Q^{-1}(\varepsilon)$ and $Q^{-1}(x)$ the inverse Q-function. Finally, since

$$u = \frac{N_0 \delta_F (2^{\frac{R_{k,i,g}}{\delta_F}} - 1)}{\rho_{k,i,g}} \approx \frac{\zeta_{k,i,g}(\varepsilon)^2 \Omega_{k,i,g}}{2(\kappa_{k,i,g} + 1)},$$

the expression of the achievable data rate can be derived as

$$R_{k,i,g} = \delta_F \log_2 \left(1 + \underbrace{\frac{\rho_{k,i,g} \zeta_{k,i,g}(\varepsilon)^2 \Omega_{k,i,g}}{2(\kappa_{k,i,g} + 1) N_0 \delta_F}}_{\text{SNR}_{k,i,g}} \right), \quad (26)$$

with $\text{SNR}_{k,i,g}$ denoting the SNR.

In the following the scaling behavior of the SNR and other properties are analyzed.

Theorem 2. *The SNR of the g -th GN scales by a factor $\sum_{u=1}^U \sum_{p_u=1}^{P_u} E_{u,p_u} \leq \xi \leq (\sum_{u=1}^U \sum_{p_u=1}^{P_u} E_{u,p_u})^2$.*

Proof. The maximum squared LoS component is obtained when the swarm is perfectly aligned with BS and GNs, which $\forall k, i, g$ and $\forall u \geq u'$ corresponds to

$$\begin{cases} \text{mod} \left(\frac{2\pi f_i}{c} (\varphi_{k,i,g,u} - \varphi_{k,i,g,u'}), 2\pi \right) = 0, & (27) \\ \text{mod} \left(\varphi_{k,i,g,u} + 2\pi f_i \frac{d_g^{\text{BRG}}}{c}, 2\pi \right) = 0, & (28) \end{cases}$$

thus allowing to recast it in the more compact notation

$$\nu_{k,i,g}^2 = \left(\sum_{u=1}^U \sum_{p_u=1}^{P_u} |\chi_{k,i,g,u,p_u}| + \lambda_g \bar{\kappa}_g^{\text{BRG}} \right)^2. \quad (29)$$

Therefore, assuming that all the patches of the swarm perfectly reflect the incident signal towards the desired GN yields

$$\begin{aligned} \nu_{k,i,g}^{\text{MAX}^2} &= \lim_{\substack{\{\psi_{k,g,u,p_u}^x\} \rightarrow 2\pi b \\ \{\psi_{k,g,u,p_u}^y\} \rightarrow 2\pi c}} \nu_{k,i,g}^2 \\ &= \left(\sum_{u=1}^U \eta_{k,g,u} \bar{\kappa}_{k,g,u}^{\text{BRG}} \sum_{p_u=1}^{P_u} a_{k,u,p_u} E_{u,p_u} + \lambda_g \bar{\kappa}_g^{\text{BRG}} \right)^2, \end{aligned}$$

for any $b, c \in \mathbb{Z}$, which is verified when

$$\begin{aligned} \phi_{k,u,p_u}^X &= -\sin \theta_{k,u}^{\text{BR}} \cos \varphi_{k,u}^{\text{BR}} - \sin \theta_{k,g,u}^{\text{RG}} \cos \varphi_{k,g,u}^{\text{RG}}, \\ \phi_{k,u,p_u}^Y &= -\sin \theta_{k,u}^{\text{BR}} \sin \varphi_{k,u}^{\text{BR}} - \sin \theta_{k,g,u}^{\text{RG}} \sin \varphi_{k,g,u}^{\text{RG}}. \end{aligned} \quad (30)$$

Consequently,

$$\lim_{\{E_{u,p_u}\} \rightarrow +\infty} \nu_{k,i,g}^{\text{MAX}^2} = \mathcal{O}(E^2),$$

being $E = \sum_{u=1}^U \sum_{p_u=1}^{P_u} E_{u,p_u}$. Furthermore, recalling the definition of the average power in (22) and the NLoS component in (24), as $\{E_{u,p_u}\} \rightarrow +\infty$ the following inequalities hold true

$$\mathcal{O}(E) \leq \Omega_{k,i,g} \leq \mathcal{O}(E^2) \quad (31)$$

and since

$$\lim_{\{E_{u,p_u}\} \rightarrow +\infty} \frac{\zeta_{k,i,g}(\varepsilon)^2}{2(\kappa_{k,i,g} + 1)} \approx \lim_{\{E_{u,p_u}\} \rightarrow +\infty} \frac{2\kappa_{k,i,g}}{2(\kappa_{k,i,g} + 1)} = \mathcal{O}(1)$$

the SNR derived in (26) scales as $\sum_{u=1}^U \sum_{p_u=1}^{P_u} E_{u,p_u} \leq \xi \leq (\sum_{u=1}^U \sum_{p_u=1}^{P_u} E_{u,p_u})^2$. ■

Corollary 2. *Following the same rationale of Theorem 2, it is not difficult to show that in case of a single drone ($U = 1$), since the alignment condition (27) is inherently satisfied, the SNR measured at a specific GN illuminated by $\bar{P} \leq P_u$ scales as $\zeta \geq \left(\sum_{p_u=1}^{\bar{P}} E_p \right)^2$.*

Corollary 3. *In the special case of no direct BS-GN link ($\lambda_g \rightarrow 0$), a single drone ($U = 1$) illuminating one GN ($G = 1$) with the entire IRS ($P = 1$), the channel follows a Rician distribution characterized by $\kappa_{k,i} = \frac{\bar{\kappa}_k^{\text{BRG}}}{\bar{\kappa}_k^{\text{BRG}}} E$ and $\Omega_k = a_k^2 \eta_k^2 (\bar{\kappa}_k^{\text{BRG}} E^2 + \tilde{\kappa}_k^{\text{BRG}} E)$, where g, u, p are omitted for brevity. Further, the corresponding SNR scales as $\zeta = E^2$.*

IV. TRAJECTORY AND PHASE SHIFT MATRIX OPTIMIZATION

Leveraging the model derived in the previous Section, an optimization problem aiming at maximizing the sum-rate of nodes and, at the same time, at fairly distributing the resources to the GNs is formulated and a resolution approach provided.

A. Problem formulation

Let assume that (i) the UAVs fly at a constant altitude $z_{k,u}^U = H \forall k, u$, (ii) the GNs have a dedicated subcarrier, and (iii) the drones have the same number of patches, i.e., $P_u = P \forall u$. Moreover, it is assumed that the UAVs have enough energy to fulfill the mission, since its duration is limited. The objective is to maximize the sum-rate $\sum_{k,i,g} R_{k,i,g}$, with $R_{k,i,g}$ given in (26) and, at the same time, to minimize the sum of the absolute differences $\sum_{g>g'} \left| \sum_{k=1}^K \sum_{i=1}^I R_{k,i,g} - R_{k,i,g'} \right|$ among these amounts, so as to provide an adequate level of fairness among users. This multi-objective problem requires the optimization of the trajectories $\mathbf{Q} = \{\mathbf{q}_{k,u}, \forall k, u\}$, the speed profiles $\mathbf{V} = \{\mathbf{v}_{k,u}, \forall k, u\}$, the acceleration profiles $\mathbf{A} = \{\mathbf{a}_{k,u}, \forall k, u\}$, and the phase shift matrices $\Phi_k^x = \{\phi_{k,u,p}^x, \forall u, p\}$ and $\Phi_k^y = \{\phi_{k,u,p}^y, \forall u, p\}$. Therefore, it can be formulated as follows

$$\max_{\mathbf{Q}, \mathbf{V}, \mathbf{A}, \{\Phi_k^x\}, \{\Phi_k^y\}} \frac{\sum_{k=1}^K \sum_{i=1}^I \sum_{g=1}^G R_{k,i,g}}{1 + \varpi \sum_{g>g'} \left| \sum_{k=1}^K \sum_{i=1}^I R_{k,i,g} - R_{k,i,g'} \right|} \quad \text{s.t.} \quad (32)$$

$$\mathbf{q}_{1,u} = \mathbf{q}_{0,u}, \quad u = 1, \dots, U, \quad (32a)$$

$$\mathbf{q}_{k+1,u} = \mathbf{q}_{k,u} + \mathbf{v}_{k,u} \delta_t + \frac{1}{2} \mathbf{a}_{k,u} \delta_t^2, \quad k = 1, \dots, K, \quad u = 1, \dots, U, \quad (32b)$$

$$\mathbf{v}_{k+1,u} = \mathbf{v}_{k,u} + \mathbf{a}_{k,u} \delta_t, \quad k = 1, \dots, K, \quad u = 1, \dots, U, \quad (32c)$$

$$\|\mathbf{v}_{k,u}\| \leq v^{\text{MAX}}, \quad k = 1, \dots, K, \quad u = 1, \dots, U, \quad (32d)$$

$$\|\mathbf{a}_{k,u}\| \leq a^{\text{MAX}}, \quad k = 1, \dots, K, \quad u = 1, \dots, U, \quad (32e)$$

$$\|\mathbf{q}_{k,u} - \mathbf{q}_{k,u'}\| \geq d^{\text{MIN}}, \quad \begin{matrix} k = 1, \dots, K, \\ \{u, u'\} = 1, \dots, U \wedge u \neq u' \end{matrix} \quad (32f)$$

where ϖ is a penalty factor that determines the fairness of the solution. The constraint (32b) and (32c) encode the possible 2D movement of the drones and (32a) describes the start point of the trajectory. Constraints (32d) and (32e) state that the speed and acceleration norm is upper-bounded by v^{MAX} and a^{MAX} , respectively. (32f) guarantees that, throughout the mission, a minimum security distance d^{MIN} is maintained among UAVs. Since the GNs positions are known, it is possible to derive the optimal patch phase shift through (30). Therefore, the problem can be rearranged in terms of scheduling, i.e., patch-GN assignment, by means of a matrix $\mathbf{S}_k = \{\mathbf{S}_{k,u,p} \forall k, u\} \in \{1, \dots, G\}^{U \times P}$. Let define an auxiliary function

$$\Theta : \mathbf{S}_k \rightarrow [\Phi_k^x, \Phi_k^y], \quad (33)$$

which takes as input the scheduling matrix and returns the optimal phases vectors. Therefore, the expression of the data rate

can be rearranged as $\overline{R}_{k,i,g} \triangleq R_{k,i,g}(\cdot, \Theta)$. The optimization problem is cast as

$$\max_{\mathbf{Q}, \mathbf{V}, \mathbf{A}, \{\mathbf{S}_k\}} \frac{\sum_{k=1}^K \sum_{i=1}^I \sum_{g=1}^G \overline{R}_{k,i,g}}{1 + \varpi \sum_{g>g'} \left| \sum_{k=1}^K \sum_{i=1}^I \overline{R}_{k,i,g} - \overline{R}_{k,i,g'} \right|} \quad (34)$$

s.t. (32a) – (32f)

Problem (34) is a MINLP problem and hence very challenging to solve. Common techniques as Block Coordinate Descent (BCD) and Successive Convex Approximation (SCA) cannot be applied in this case, since the channel model involves cosine interference patterns, which are not convexifiable. Therefore, in this work a DRL approach is employed to achieve a quasi-optimal solution.

B. Background on Proximal Policy Optimization (PPO)

DRL is a branch of machine learning which involves an agent interacting via actions with an environment. Among the vast plethora of available DRL algorithms, PPO [25] represents the most adopted solution due to the noticeable performance. The benefits brought by PPO approach are numerous. The most important ones are (i) ease of implementation, (ii) low complexity, (iii) sample efficiency, and (iv) few hyperparameters are needed in the learning process. It ensures a smoother training, compared with other approaches, by constraining the new policy to not excessively differ from the previous one. This yields lower variance in the training process and prevents the agent from taking unrecoverable paths. In this work, it is adopted to solve problem (34) and hence achieve a quasi-optimal solution. PPO uses an Actor-Critic (AC) approach, which employs two Deep Neural Networks. The Actor, denoted as $Q_{\vartheta}(s_k, a_k)$, at each discrete time instant k , observes the current state s_k . Based on that, it takes a certain action a_k , according to the current policy $\pi_{\vartheta}(s_k|a_k)$, being ϑ the corresponding parameter set. Consequently, it receives a reward r_k and moves to the next state s_{k+1} . The Critic, denoted as $V_{\tau}(s_k)$, evaluates the action taken by the Actor and provides a rating. Based on this value, the Actor improves the current policy π_{ϑ} , adopting a gradient descent algorithm, to take better, or to avoid worst, actions in the future. Concretely, the aim of this framework is to derive the optimal policy π_{ϑ}^* from the transition tuples $\langle s_k, a_k, r_k, s_{k+1} \rangle$ that maximize the discounted cumulative sum of all future rewards. Hence, $\pi_{\vartheta}(s_k|a_k)$ maps each state to the probability $p(s_{k+1}|s_k, a_k)$ of taking action a_k .

PPO offers several advantages [25] over classical AC algorithm. Specifically, it incorporates a trust region optimization approach, ensuring stable policy updates and preventing catastrophic policy changes. The clipped surrogate objective further constrains the policy update range, enhancing robustness. Moreover, the multiple epochs of mini-batch updates allow efficient utilization of collected data, thus improving learning effectiveness. Differently from the AC algorithm, PPO demonstrates scalability, sample efficiency, and compatibility with both continuous and discrete action spaces. It strikes

a balance between exploration and exploitation, facilitating optimal policy learning. In contrast, classical AC algorithms often lack such robustness, require careful hyperparameter tuning, exhibit higher variance in updates, and may have limited exploration capabilities.

To provide more technical background, the rationale behind the Policy Gradient Methods, and hence of the PPO algorithm, are hereby discussed.

The algorithm compares the current and new policies to maximize the following objective function

$$\mathcal{L}_{\vartheta, \vartheta'}(s_k, a_k) = \min \left(\pi_{\vartheta, \vartheta'}(s_k, a_k) \hat{A}_k^{\pi}(s_k, a_k), \right. \\ \left. \text{clip} \left(\pi_{\vartheta, \vartheta'}(s_k, a_k), 1 - \epsilon, 1 + \epsilon \right) \hat{A}_k^{\pi}(s_k, a_k) \right)$$

with

$$\pi_{\vartheta, \vartheta'}(s_k, a_k) = \frac{\pi_{\vartheta}(s_k|a_k)}{\pi_{\vartheta'}(s_k|a_k)}$$

being the probability ratio, ϑ' the old policy parameters, and ϵ a small constant. Moreover, $A_k^{\pi}(\cdot, \cdot)$ is the advantage function defined as:

$$\hat{A}_k^{\pi}(s_k, a_k) = (1 - m) \sum_{l'=k}^K m^{l'-k} A_{k,l'}^{\pi}(s_k, a_k), \quad (35)$$

$$A_{k,l'}^{\pi}(s_k, a_k) = \sum_{l=k}^{l'-1} \zeta^{l-k} r_{k+l} - V_{\tau}(s_k) + \zeta^k V_{\tau}(s_{k+l}), \quad (36)$$

where $m \in [0, 1]$ is a smoothing factor and $A_{k,l'}^{\pi}(s_k, a_k)$ is the l' -step look-ahead advantage function, with $l' = k, \dots, K$. $\text{clip}(\cdot, \cdot, \cdot)$ denotes the clipping function [25] which bounds the first argument in range of the last two, such that the new policy is not allowed to go too far from the old. While, in case of positive advantage the objective reduces to

$$\mathcal{L}_{\vartheta, \vartheta'}(s_k, a_k) = \min \left(\pi_{\vartheta, \vartheta'}(s_k, a_k), 1 + \epsilon \right) \hat{A}_k^{\pi}(s_k, a_k), \quad (37)$$

where the first term limits to how much the objective can increase. Similarly, for negative advantage:

$$\mathcal{L}_{\vartheta, \vartheta'}(s_k, a_k) = \max \left(\pi_{\vartheta, \vartheta'}(s_k, a_k), 1 - \epsilon \right) \hat{A}_k^{\pi}(s_k, a_k). \quad (38)$$

Therefore, the clipping function provides a regularization of the objective function by preventing extreme changes, and the hyperparameter ϵ represents how far the new policy $\pi_{\vartheta}(s_k|a_k)$ can go from the old one $\pi_{\vartheta'}(s_k|a_k)$. Finally, the two networks are trained using a stochastic gradient descent method, namely Adam [26], over a mini-batch of D samples. For each one, the parameters of the policy $\pi_{\vartheta}(s_k|a_k)$ and the Actor network are updated as

$$\vartheta^* = \arg \max_{\vartheta} \mathcal{L}_{\vartheta, \vartheta'}(s_k, a_k), \quad (39)$$

while the Critic parameters as

$$\tau^* = \arg \min_{\tau} \left(\hat{A}_k^{\pi}(s_k, a_k) \right)^2. \quad (40)$$

For further details the reader is referred to [25] and [27].

C. Proposed Solution

Problem (34) is reformulated as a MDP, which is a mathematical framework at the foundation of DRL. In this way, it is possible to embed the optimization problem into a PPO-based algorithm.

First, to increase training speed and prevent divergence, the scheduling matrix is normalized such that $-1 \leq \mathbf{S}_{k,u,p} \leq 1$, as follows

$$\bar{\mathbf{S}}_{k,u,p} = \left\lceil G \frac{\mathbf{S}_{k,u,p} + 1}{2} \right\rceil. \quad (41)$$

Similarly, also the acceleration profile vectors $\mathbf{a}_{k,u}$ are normalized as $\bar{\mathbf{a}}_{k,u} = \frac{\mathbf{a}_{k,u}}{a^{\text{MAX}}}$. The reference MDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \varsigma \rangle$, is described by:

- *State Space* \mathcal{S} : The set of all the possible states $s_k \in \mathcal{S}$ that can be observed by the agent while interacting with the environment with

$$s_k = \{ \mathbf{q}_{k,1}^\top, \dots, \mathbf{q}_{1,U}^\top, \dots, \mathbf{S}_{k,1,1}, \dots, \mathbf{S}_{k,1,P}, \dots, \mathbf{S}_{k,U,1}, \dots, \mathbf{S}_{k,U,P} \} \quad (42)$$

- *Action Space* \mathcal{A} : The set of all possible actions $a_k \in \mathcal{A}$ that the agent can perform during a timestep where

$$a_k = \{ \bar{\mathbf{a}}_{k,1}^\top, \dots, \bar{\mathbf{a}}_{k,U}^\top, \dots, \bar{\mathbf{S}}_{k,1,1}, \dots, \bar{\mathbf{S}}_{k,1,P}, \dots, \bar{\mathbf{S}}_{k,U,1}, \dots, \bar{\mathbf{S}}_{k,U,P} \} \quad (43)$$

- *Transition Probability* \mathcal{P} : The set of probabilities $p_k \in \mathcal{P}$ which denote the transition from a state s_k to s_{k+1} .
- *Reward function* \mathcal{R} : The rewards $r_k \in \mathcal{R}$ obtained by maximizing the objective function. In this case, it has been already defined in problem (34).
- *Discount factor* ς : The discount factor that determines the importance of future rewards, where $0 \leq \varsigma \leq 1$.

At the beginning, it is necessary to initialize the structures needed for the computation. In particular, the initial positions of the drones is set according to (32a). After the initialization phase, at each timeslot k , the agent observes from the environment the current positions of the drones \mathbf{Q}_k and the assignment of the patches \mathbf{S}_k . Then, the agent provides the acceleration profile \mathbf{A}_k and the new scheduling matrix $\bar{\mathbf{S}}_k$ in input to the algorithm that computes the new positions and speed profiles of the drones, according to (32b) and (32c), and the data rate of each GNs. The final output of the computation is the new state (observation), the generated reward, and a flag that notifies whether the mission has terminated or not. Clearly, whereas the mission ends before K , the cumulative reward would be lower; this mechanism is employed to satisfy the remaining constraints (32d),(32e), and (32f).

V. NUMERICAL RESULTS AND DISCUSSION

In this Section a simulation campaign is carried out to investigate different aspects of the proposed model.

In particular, the analysis first focuses on the validation of the channel model from a probabilistic standpoint by comparing the derived approximation with Monte Carlo simulations. Then, the radiation pattern corresponding to different patch configurations is studied. Moreover, results are shown for the optimization problem defined in Section IV, to demonstrate the

Symbol	Value	Symbol	Value
K	{1, 10}	G	{1, 2, 20}
U	{1, 3}	I	{1, 2, 20}
w	10^{-4} [m ²]	$\rho_{k,i,g}, \forall k, i, g$	20 [dBm]
$d^x = d^y$	10^{-2} [m]	N_0	-174 [dBm/Hz]
f_c	10.5 [GHz]	κ^{MIN}	6 [dB]
δ_f	10 [kHz]	κ^{MAX}	10 [dB]
α^{BG}	{3,4}	$a_{k,u,p,u}, \forall k, u, p$	1
δ_t	1 [s]	ϖ	5e-4, 1.25e-4
ε	0.01	d^{MIN}	5 [m]
v^{MAX}	30 [m/s]	ς	0.3
a^{MAX}	5 [m/s ²]	ϵ	0.2

TABLE II: Parameter set used in the simulations.

potential of the diversity introduced by the patches. Finally, the solution is compared with the case of absence of IRS-equipped drones and with a baseline approach implementing a random patch scheduling.

In all the simulations, all the involved antennas are isotropic, thus $F^{\text{BRG}}(\cdot, \cdot) = F^{\text{BRG}}(\cdot, \cdot, \cdot, \cdot) = 1$, with unitary gains. All the considered IRS are composed by $N \times M = 48 \times 48$ elements. Moreover, according to [19], the cascaded channel BS-UAV-GN is characterized by $\beta^{\text{BRG}} = \frac{c^2 w}{64\pi^3 f_i^2}$ and $\alpha^{\text{BR}} = \alpha^{\text{RG}} = 2$, while the direct weak BS-GN link by $\beta^{\text{BG}} = \frac{c^2}{16\pi^2 f_g^2}$ and $\alpha^{\text{BG}} = 4$. Besides, it is assumed $\kappa^{\text{MIN}} = \kappa^{\text{MIN,BG}} = \kappa^{\text{MIN,RG}} = \kappa^{\text{MIN,BR}} = \kappa^{\text{MIN,BR}}$ and $\kappa^{\text{MAX}} = \kappa^{\text{MAX,BG}} = \kappa^{\text{MAX,RG}} = \kappa^{\text{MAX,BR}} = \kappa^{\text{MAX,BR}}$, for all k, g, u . The main considered parameters, unless otherwise specified, are summarized in Table II.

A. Channel Model Analysis

Without loss of generality, the first scenario considers $G = 1$ GN, located at $\mathbf{q}_1^g = [50 \ 0 \ 0]^\top$, served by a BS at $\mathbf{q}^{\text{BS}} = [-50 \ 0 \ 0]^\top$. The communication is assisted by a drone, positioned at $\mathbf{q}_{1,1} = [0 \ 0 \ 50]^\top$, equipped with just $P = 1$ patch, i.e., the whole IRS. To provide a comprehensive analysis, four cases are considered: (i) absence of the direct BS-GN link ($\lambda_g = 0$), (ii) scattered (Rayleigh) BS-GN link ($\kappa_g^{\text{BG}} = 0$), (iii) worst possible UAV-BS alignment (Worst case), and (iv) best possible UAV-BS alignment (Best case). Moreover, to assess the performance of the investigated scenarios, define the normalized SNR $\forall k, i, g$, as

$$\overline{\text{SNR}}_{k,i,g} = \frac{N_0 \delta_f}{\rho_{k,i,g}} \text{SNR}_{k,i,g}.$$

It is worth noting that the data rate derived in equation (26) is directly proportional to the normalized SNR when the latter is expressed in decibels (due to the presence of the logarithm), as in the graphs shown in this Section. Figure 2 represents the comparison between the proposed model and a Monte Carlo simulation for $5 \cdot 10^6$ realizations. As a matter of fact, the proposed model provides an accurate approximation in terms of PDF and $\overline{\text{SNR}}$ curves. Moreover, it can be observed that, even if the BS-GN link exists, the UAV-BS alignment is crucial: in the worst case scenario the channel quality results to be deteriorated with respect to the pure reflection case ($\lambda_g = 0$). Figure 3 corroborates the theoretical results obtained in Corollary 2. Indeed, a large number of PRUs implies a quadratic scale of the $\overline{\text{SNR}}$. Clearly, in the four

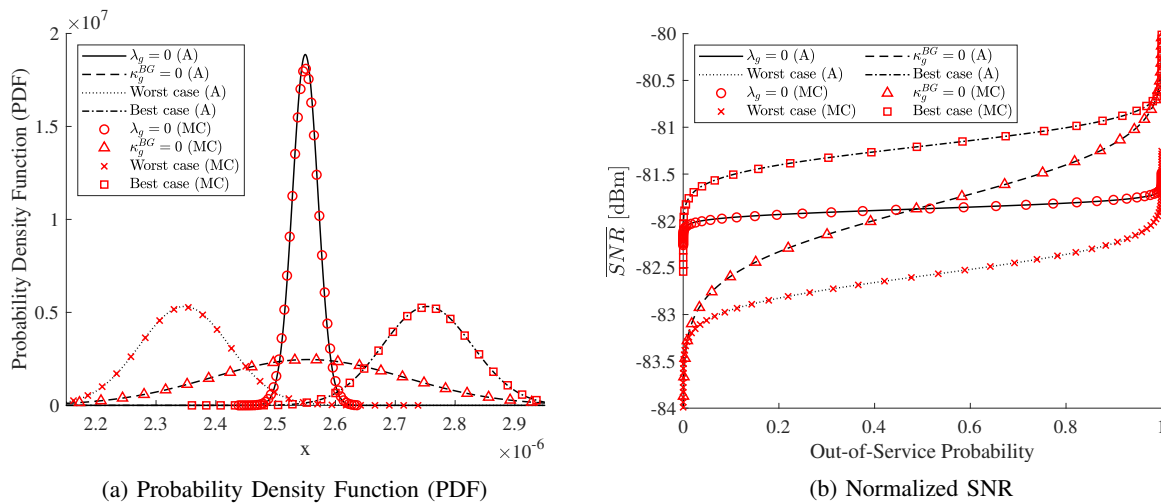


Fig. 2: (A)approximation vs Monte Carlo (MC).

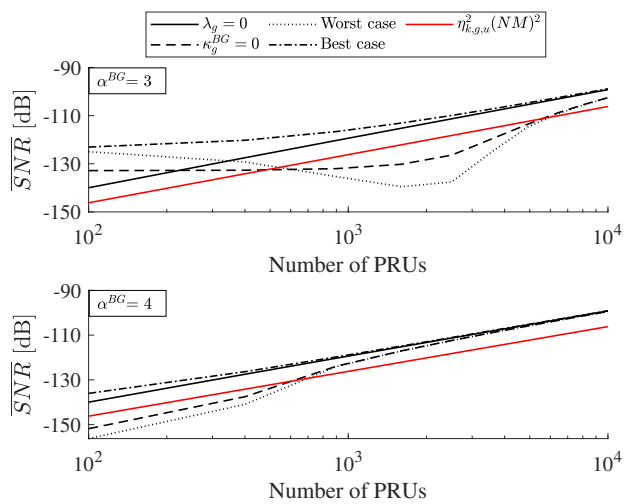


Fig. 3: $\overline{\text{SNR}}$ with respect to number of PRUs.

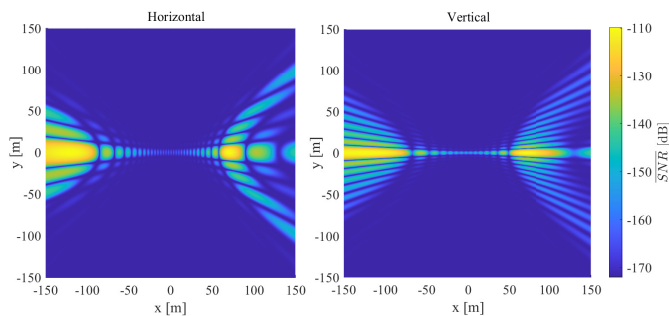


Fig. 4: $\overline{\text{SNR}}$ for $P_u = 2$ patches divided horizontally (left) and vertically (right).

considered cases, the convergence depends on the involved system setup and environmental surrounding conditions. It is worth noting that a reduced path loss exponent $\alpha^{BG} = 3$ produces, in the Best case and for a low number of elements, an higher gain with respect to the pure reflection case ($\lambda_g = 0$). When a higher path loss coefficient is considered, i.e., $\alpha^{BG} = 4$,

this gap significantly reduces since the direct link brings a minor contribution to the SNR.

The second scenario involves the same UAV and IRS, but considers $G = 2$ GNs located at $\mathbf{q}_1^g = [-120 \ 0 \ 0]^T$ and $\mathbf{q}_2^g = [70 \ 0 \ 0]^T$, connected to a BS at $\mathbf{q}^{\text{BS}} = [-50 \ 0 \ 0]^T$ with $\lambda_g = 0$. Two cases are investigated. In the first one, the IRS surface is divided horizontally, while in the second one, vertically. Then, the phase shift matrices are set so that each user has its own patch assigned. From Figure 4 it clearly emerges that the produced radiation patterns yield a different $\overline{\text{SNR}}$ perceived in the environment in the two cases. Indeed, the side of the patch characterized by more elements leads to a narrower beam in the orthogonal direction and viceversa. This phenomenon, which can be observed especially in horizontally-divided case, is also influenced by the UAV-GN distance. Therefore, the shape design is particularly important since remarkably modifies the irradiated area, and has different implications based on the use case considered. For instance, if the objective is to maximize channel quality perceived by multiple GNs, the neighborhood of a illuminated user benefits from configurations that produces a larger fingerprint. At the same time, if the aim is to establish secure communications, patches must be designed on the opposite criteria, thus hinder eavesdropping. An agnostic option is to split the IRS into square-shaped patches that can be directed towards specific GNs, thus implying a customizable radiation pattern which guarantees, however, a computation complexity much lower than solution proposed in [18].

B. Results on trajectory and phase shift optimization

It is considered a scenario in which $U = 3$ drones, starting from $\mathbf{q}_{1,u} = [0 \ -150 \ 50]^T, \forall u$, are in charge of fairly serving $G = 20$ nodes randomly deployed on the ground. The BS is located at $\mathbf{q}^{\text{BS}} = [0 \ 0 \ 0]^T$. The Actor and Critic neural networks of the PPO algorithm are characterized by four hidden layers of 1024 neurons each, trained for $3 \cdot 10^5$ epochs.

First, it is considered the case in which each drone is equipped with a IRS split into $P_u = 9 \ \forall u$ patches, by adopting

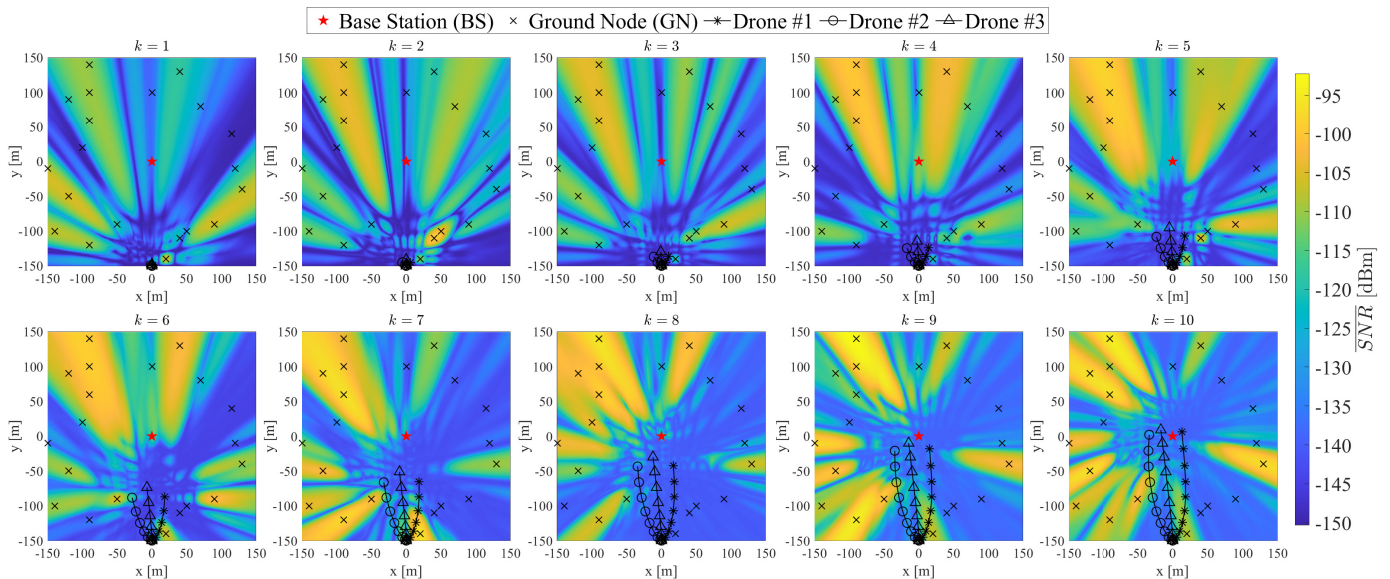


Fig. 5: Drones' trajectories and normalized SNR for $G = 20$, $U = 3$, $P_u = 9$, and $\varpi = 1.25 \cdot 10^{-4}$.

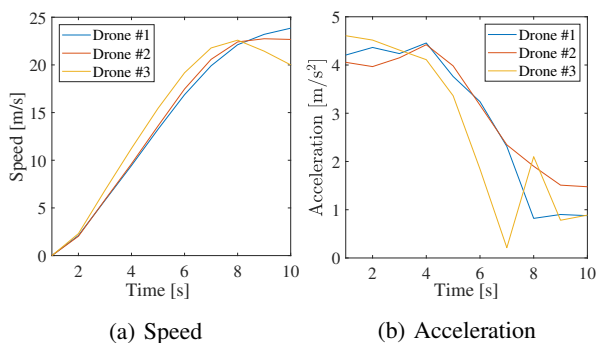


Fig. 6: Speed and acceleration for $P_u = 4$ and $\varpi = 1.25 \cdot 10^{-4}$.

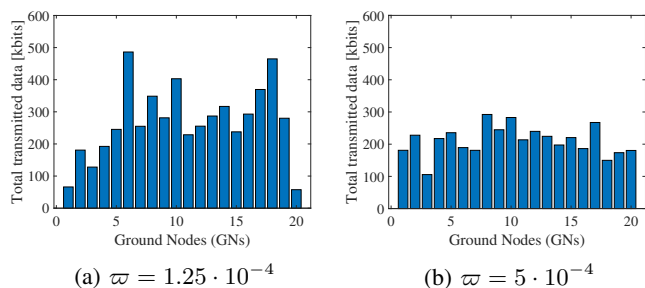


Fig. 7: Total amount of data received by nodes, with $P_u = 4$.

a low fairness factor $\varpi = 1.25 \cdot 10^{-4}$. Figure 5 shows the trajectories of the drones during the mission, with heatmap representing the values of the normalized SNR considering only the BS-UAV-GN channel. It is visible that, during the mission, the drones serve different regions at each instant, with the aim of maximizing the sum-rate while providing a ϖ -dependent allocation to the users. The swarm flies, as result of the optimization, from the starting location towards the center of the area, where the BS is located. Consequently, the patches of the different UAVs are allocated to provide an optimal

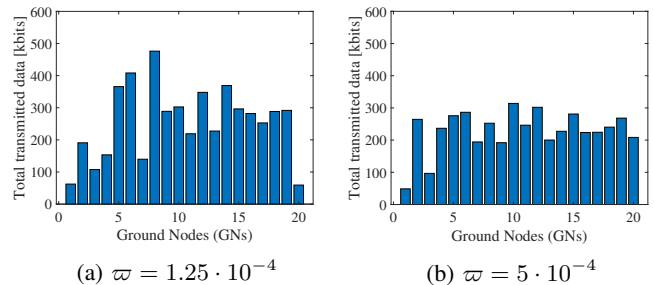


Fig. 8: Total amount of data received by nodes, with $P_u = 9$.

coverage to all users, thus improving their SNRs. Moreover, in Figure 6, the speed and acceleration profiles, i.e., $\|\mathbf{v}_{k,u}\|$ and $\|\mathbf{a}_{k,u}\|$, are shown for each drone over time. According to (32b) and (32c), the drones accelerate yielding an increase in the speed, without violating the bound constraints (32d) and (32e). This proves that the strategy adopted and discussed in Section IV-C allows the PPO-based algorithm to learn from tuples that satisfy the imposed limits.

To give more insights about the exchanged data volume and to validate the effectiveness of the proposed algorithm in other configurations, the IRS are split into $P_u = 4$ and $P_u = 9$ patches. Moreover, two values for the fairness parameter ϖ are considered, i.e., $1.25 \cdot 10^{-4}$ and $5 \cdot 10^{-4}$. The first value penalizes the objective function more, thus obtaining a lower fairness with respect to the second one, which implies a higher balance in the resource distribution.

In Figure 7, it is depicted the cumulative amount of data transmitted from each GNs to the BS through the direct link plus the reflected one, for $P_u = 4$. In case of higher fairness factor, i.e., $\varpi = 1.25 \cdot 10^{-4}$, the total amount of received data is 5.37 Mbits with a coefficient of variation (ratio between standard deviation and mean) of 0.41. On the contrary, in the second case, i.e., $\varpi = 5 \cdot 10^{-4}$, a lower amount of data is received by the BS, i.e., 4.21 Mbits, but with a coefficient of

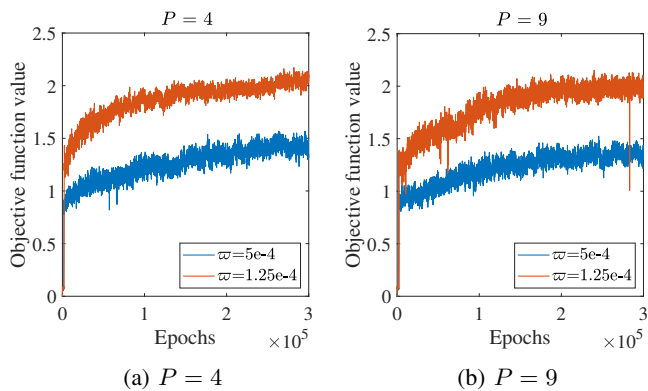


Fig. 9: Convergence curve of the training process.

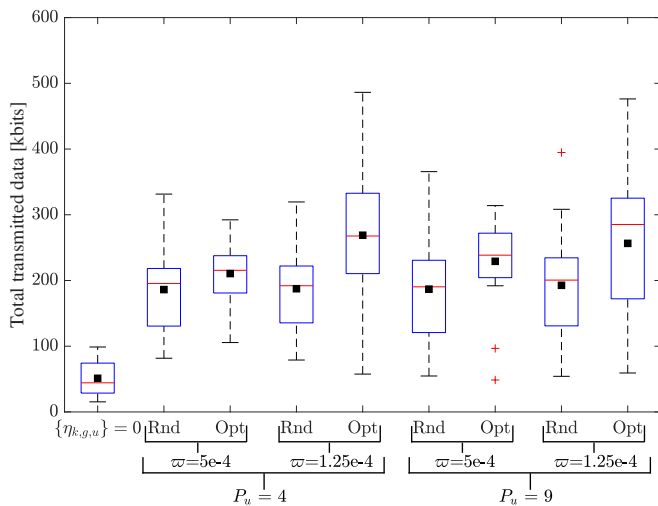


Fig. 10: Comparison among total data transmitted amounts with optimized trajectories for random (Rnd) and optimized (Opt) scheduling. The black squares represent the mean value.

variation equal to 0.2. This implies that a trade-off between fairness and data transmission exists: it is due to the fact that in case of lower fairness the drones can focus on reflecting the signals from the GN closer to the current position, which in turns implies that farther nodes are less served.

For the same scenario parameters, a simulation with $P_u = 9$ patches (for each drone) is carried out. As expected, Figure 8 confirms the same trade-off behavior highlighted in the previous case. As a matter of fact, with $P_u = 9$ patches, in the first case a total of 5.12 Mbits have been transmitted with a coefficient of variation equal to 0.42, while in second case the values are 4.58 Mbits and 0.27, respectively.

Moreover, in Figure 9 it can be seen the convergence of the neural network training process in all the configurations discussed above. Finally, Figure 10 summarizes the results and the discussion above by comparing the obtained values with the case in which the swarm is absent and only the BS serves the GNs, i.e., $\{\eta_{k,g,u}\} = 0 \forall k, g, u$. Moreover, to provide a benchmark, the proposed solution is also compared with a baseline approach, in which the optimized trajectories are maintained but a random patch scheduling is implemented ($5 \cdot 10^5$ trials have been performed). As expected, in all cases,

the drones provide a substantial enhancement of the channel capacity, especially when an optimized solution is employed. Indeed, the scheduling plan obtained from the PPO performs better than that related to the baseline approach. In particular, this is always true in terms of average data rate. Nonetheless, when a higher number of patches is employed a higher fairness factor is required to guarantee a uniform resource distribution.

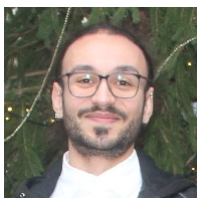
VI. CONCLUSIONS

In this work, a comprehensive channel model for UAV-aided IRS-assisted OFDMA communications has been derived. Differently from other contributions, it considers the presence of a swarm of drones equipped with IRSs, which are split into an arbitrary number of patches to simultaneously serve multiple GNs. The proposed model inherently captures the constructive/destructive interference among users which can also experience different propagation conditions, due to the adoption of Rice fading. Relying on this channel model, a realistic communication scenario has been investigated, which led to the formulation of a Multi-Objective MINLP problem, aiming to fairly distribute the resources among nodes, i.e., to maximize the sum-rate of GNs and, at the same time, to minimize the differences among rates. This required the joint optimization of the phase shift matrices of the IRSs and the trajectories of the drones. To overcome the intractability of the non-convex cosine patterns related to wave interference, the problem has been rearranged as a MDP and solved via a DRL method, namely PPO. The validity of this work has been corroborated by extensive simulations, for different parameter settings, including the comparison with a baseline approach. Results demonstrated that (i) channel capacity benefits from IRS employment and that (ii) the proposed solution outperforms the baseline in terms of total and average transmitted data. Future work includes adding subcarrier scheduling in the joint optimization and considering the presence of multiple BSs. Moreover, models that capture the dependency of amplitude and phase of the IRS elements will be considered.

REFERENCES

- [1] M. Gharibi, R. Boutaba, and S. L. Waslander, "Internet of drones," *IEEE Access*, vol. 4, pp. 1148–1162, 2016.
- [2] P. Boccadoro, D. Striccoli, and L. A. Grieco, "An extensive survey on the Internet of Drones," *Ad Hoc Netw.*, vol. 122, p. 102600, 2021.
- [3] L. Chettri and R. Bera, "A Comprehensive Survey on Internet of Things (IoT) Toward 5G Wireless Systems," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16–32, 2020.
- [4] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, "6G Internet of Things: A Comprehensive Survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, 2022.
- [5] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, Wide-Area, and Non-Terrestrial IoT: A Survey on 5G Advances and the Road Toward 6G," *IEEE Commun. Surv. Tutor.*, vol. 24, no. 2, pp. 1117–1174, 2022.
- [6] Z. Wei, M. Zhu, N. Zhang, L. Wang, Y. Zou, Z. Meng, H. Wu, and Z. Feng, "UAV-Assisted Data Collection for Internet of Things: A Survey," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15460–15483, 2022.
- [7] M. Abrar, U. Ajmal, Z. M. Almohaimeed, X. Gui, R. Akram, and R. Masroor, "Energy Efficient UAV-Enabled Mobile Edge Computing for IoT Devices: A Review," *IEEE Access*, vol. 9, pp. 127 779–127 798, 2021.
- [8] R. Liu, Q. Wu, M. Di Renzo, and Y. Yuan, "A Path to Smart Radio Environments: An Industrial Viewpoint on Reconfigurable Intelligent Surfaces," *IEEE Wirel. Commun.*, pp. 1–7, 2022.

- [9] R. Chen, M. Liu, Y. Hui, N. Cheng, and J. Li, "Reconfigurable Intelligent Surfaces for 6G IoT Wireless Positioning: A Contemporary Survey," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23 570–23 582, 2022.
- [10] A. C. Pogaku, D.-T. Do, B. M. Lee, and N. D. Nguyen, "UAV-Assisted RIS for Future Wireless Communications: A Survey on Optimization and Performance Analysis," *IEEE Access*, vol. 10, pp. 16 320–16 336, 2022.
- [11] K. Tekbryik, G. Karabulut Kurt, A. R. Ekti, and H. Yanikomeroglu, "Reconfigurable Intelligent Surfaces in Action: For Nonterrestrial Networks: Employing Reconfigurable Intelligent Surfaces," *IEEE Veh. Technol. Mag.*, pp. 2–10, 2022.
- [12] A. S. Abdalla, T. F. Rahman, and V. Marojevic, "UAVs with Reconfigurable Intelligent Surfaces: Applications, Challenges, and Opportunities," 2020. [Online]. Available: <https://arxiv.org/abs/2012.04775>
- [13] B. Shang, R. Shafin, and L. Liu, "UAV Swarm-Enabled Aerial Reconfigurable Intelligent Surface," 2021. [Online]. Available: <https://arxiv.org/abs/2103.06361>
- [14] Z. Wei, Y. Cai, Z. Sun, D. W. K. Ng, J. Yuan, M. Zhou, and L. Sun, "Sum-Rate Maximization for IRS-Assisted UAV OFDMA Communication Systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2530–2550, 2021.
- [15] Q. Tao, J. Wang, and C. Zhong, "Performance Analysis of Intelligent Reflecting Surface Aided Communication Systems," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2464–2468, 2020.
- [16] A. M. Salhab and M. H. Samuh, "Accurate Performance Analysis of Reconfigurable Intelligent Surfaces Over Rician Fading Channels," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1051–1055, 2021.
- [17] M. A. Al-Jarrah, E. Alsusa, A. Al-Dweik, and M.-S. Alouini, "Performance Analysis of Wireless Mesh Backhauling Using Intelligent Reflecting Surfaces," *IEEE Trans. Wireless Commun.*, 2021.
- [18] G. Iacovelli, A. Coluccia, and L. A. Grieco, "Channel Gain Lower Bound for IRS-Assisted UAV-Aided Communications," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3805–3809, 2021.
- [19] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless Communications With Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2021.
- [20] M. A. Al-Jarrah, K.-H. Park, A. Al-Dweik, and M.-S. Alouini, "Error Rate Analysis of Amplitude-Coherent Detection over Rician Fading Channels with Receiver Diversity," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 134–147, 2019.
- [21] C. You and R. Zhang, "3D Trajectory Optimization in Rician Fading for UAV-Enabled Data Harvesting," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3192–3207, 2019.
- [22] N. O'Donoghue and J. M. F. Moura, "On the Product of Independent Complex Gaussians," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1050–1063, 2012.
- [23] H. Lu, Y. Zeng, S. Jin, and R. Zhang, "Aerial Intelligent Reflecting Surface: Joint Placement and Passive Beamforming Design With 3D Beam Flattening," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4128–4143, 2021.
- [24] 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, 2017.
- [25] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017. [Online]. Available: <https://arxiv.org/abs/1707.06347>
- [26] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2017. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [27] S. Parisi, V. Tangkaratt, J. Peters, and M. E. Khan, "TD-regularized actor-critic methods," *Mach. Learn.*, vol. 108, pp. 1467–1501, 2019.



Giovanni Iacovelli received the Ph.D. in electrical and information engineering from Politecnico di Bari, Bari, Italy, in December 2022. His research interests include Internet of Drones, Machine Learning, Optimization and Telecommunications. Currently, he is an Assistant Professor in Telecommunications at the Department of Electrical and Information Engineering, Politecnico di Bari.



Angelo Coluccia received the Ph.D. degree in Information Engineering in 2011 and is currently an Associate Professor of Telecommunications at the Department of Engineering, University of Salento (Lecce, Italy). His research interests are in the area of multi-channel, multi-sensor, and multi-agent statistical signal processing for detection, estimation, localization, and learning problems. Relevant application fields are radar, wireless networks (including 5G and beyond), and emerging network contexts (including intelligent cyber-physical systems, smart devices, and social networks). He is Senior Member of IEEE, Member of the Sensor Array and Multichannel Technical Committee for the IEEE Signal Processing Society, and Member of the Technical Area Committee in Signal Processing for Multisensor Systems of EURASIP.



L. Alfredo Grieco is a full professor in telecommunications at Politecnico di Bari. His research interests include Internet of Things, Future Internet Architectures, and Nano-communications. He serves as Founder Editor in Chief of the Internet Technology Letters journal (Wiley) and as Associate Editor of the IEEE Transactions on Vehicular Technology journal (for which he has been awarded as top editor in 2012, 2017, and 2020).