

ESN Reinforcement Learning for Spectrum and Flight Control in THz-Enabled Drone Networks

Sabarish Krishna Moorthy, *Student Member, IEEE*, Maxwell McManus, *Student Member, IEEE* and Zhangyu Guan, *Senior Member, IEEE*

Abstract—Terahertz (THz)-band communications have been envisioned as a key technology to support ultra-high-data-rate applications in 5G-beyond (or 6G) wireless networks. Compared to the microwave and mmWave bands, the main challenges with the THz band are in its i) large path loss hence limited network coverage and ii) visible-light-like propagation characteristics hence poor support of mobility in blockage-rich environments. This paper studies quantitatively the applicability of THz-band communications in blockage-rich mobile environments, focusing on a new network scenario called *FlyTera*. In *FlyTera*, a set of hotspots mounted on flying drones collaboratively provide data streaming services to ground users, in the microwave, mmWave and THz bands. We first provide a mathematical formulation of the *FlyTera* control problem, where the objective is to maximize the network spectral efficiency by jointly controlling the flight of the drone hotspots, their association to the ground users, and the spectrum bands used by the users. To solve the resulting problem, which is shown to be a mixed integer nonlinear nonconvex programming (MINLP) problem, we design distributed solution algorithms based on a combination of echo state learning and reinforcement learning. An extensive simulation campaign is then conducted with SimBAG, a newly developed Simulator of Broadband Aerial-Ground wireless networks. It is shown that no single spectrum band can meet the requirements of high data rate and wide coverage simultaneously. Moreover, from the network-level point of view, THz-band communications can significantly benefit from the mobility of the flying drones, and on average 4–6 times higher (rather than lower) throughput can be achieved in mobile than in static environments.

Index Terms—Terahertz Band; Mmwave Band; Microwave Band; Drone Networks; Echo State Network; Reinforcement Learning.

I. INTRODUCTION

With the advancements of new material development and transceiver design [2]–[6], THz (i.e., Terahertz, with frequency ranging from 100 GHz to 10 THz) communications have been envisioned as a key technology to meet the increasing demands of bandwidth-hungry applications in 5G-beyond and 6G wireless networks, such as wireless virtual/augmented reality (VR/AR) [7], [8], high-data-rate communications [9], [10], vehicular networks [11], among others [12]–[14]. However, compared to lower frequency bands, e.g., sub-6 GHz and mmWave bands, there are two main challenges with the THz band. First, its communication range is significantly

reduced because of the large signal attenuation in THz band in radio in-air environments. For example, the attenuation due to water vapor and oxygen absorption is approximately 0.6 – 1000 dB/km for THz band, while it is 0.01 dB/km at sub-6 GHz band and 0.3 – 0.6 dB/km for the mmWave band [15]. Second, the THz links can be easily blocked because of the visible-light-like directional waves in extremely high frequency range.

In the past few years, significant research efforts have been directed towards addressing these challenges, focusing on either THz or mmWave bands. For example, in [16], [17] Han et al. propose a multi-wideband waveform design for the THz band, which improves the communication distance by dynamically varying the rate and the transmit power on each sub-window. The concept of ultra-massive MIMO communications was studied in [18] to increase the communication distance and the achievable capacity of THz-band communications. In mobile environments, fast beam search and alignment schemes have been proposed in [19]–[22]. For example, Hassanieh et al. propose Agile-Link in [19], which can provably find the optimal direction in logarithmic number of measurements. BeamSpy is proposed in [21] to instantaneously predict the quality of mmWave beams without the costly beam searching. Readers are referred to [23], [24] and references therein for an excellent survey of the main results in this area.

Most of these above discussed works require to redesign the lower layers (i.e., physical and link) of the communicating devices' protocol stack, and hence are not backward compatible. Moreover, these work either focus on only single link or static networking scenarios, while the applicability of THz-band communications for mobile wireless networking has not been thoroughly explored so far.

Novelty and Contributions. In this work we aim to understand from a network perspective the applicability of THz-band communications and how it can complement the lower-frequency bands in mobile blockage-rich environments. Inspired by the newly emerging drone cells [9], [25]–[28], in this paper we focus on a network scenario called *FlyTera*, where a set of hotspots mounted on flying drones collaboratively provide data streaming services to ground users in the microwave, mmWave and THz bands. In *FlyTera*, we consider UAVs endowed with multiple radio interfaces to achieve a good tradeoff between high-data-rate communications and large network coverage [29]. It is worth pointing out that it is feasible to deploy multiple radio interfaces on modern UAVs, since a mmWave or THz antenna array can be easily packed in a small area because of the short wavelength in those

A preliminary shorter version of this paper appeared in the Proceedings of IEEE International Conference on Sensing, Communication and Networking (SECON), Como, Italy, June 2020 [1].

Sabarish Krishna Moorthy, Maxwell McManus and Zhangyu Guan are with the Department of Electrical Engineering, The State University of New York at Buffalo, Buffalo, NY 14260. E-mail: {sk382, memcmanu, guan}@buffalo.edu.

frequency bands. Universal transceivers across multiple frequency bands can also be enabled by the recent advancements of material and manufacturing technologies [30]. Moreover, with the advancement of battery technologies, the flight time of UAVs has been significantly prolonged with enhanced capability of operating complex computing algorithms. An example is hybrid solar cell HEPS, a lithium-ion polymer battery and super-capacitor bank based propulsion system [31]. Additionally, several techniques have been proposed to further extend the flight time of UAVs, e.g., Simultaneous Wireless Information and Power Transmission (SWIPT) [32], automatic battery replacement mechanism (Endless Flyer) [33], and solar-cell powered UAV [34].

We consider *FlyTera* because the drone hotspots can be deployed dynamically at network run time, and hence i) it is more likely for them to establish line-of-sight (LOS) links to their users in blockage-rich environments; and ii) the network coverage and spectral efficiency can be enhanced by dynamically deploying more drone hotspots in those areas with higher user density and higher traffic demand. There are two major challenges to address in *FlyTera*. First, it is challenging for the distributed drone hotspots to coordinate with each other to achieve extended network coverage while still maintaining high-data-rate wireless links. This is because a drone hotspot may fly away from the ground network infrastructure when it moves closer to the users, hence reducing the data rate of the backhaul link. Second, the spectrum access and association strategies of the ground users are closely coupled with the drone hotspot locations and the interference levels on each spectrum band. This makes it both essential yet challenging to achieve a good tradeoff between network coverage and network spectral efficiency. To the best of our knowledge, this is the first work studying the joint access and flight control jointly considering the microwave, mmWave and THz bands in mobile blockage-rich environments. We claim the following three main contributions:

- We first formulate mathematically the control problem in *FlyTera*, where the objective is to maximize the network-wide spectral efficiency by jointly determining the flight of the drone hotspots, their association to the ground users, and the spectrum bands used by the users. It is shown that the resulting problem is a mixed integer nonlinear non-convex programming (MINLP) problem.
- We design distributed algorithms to solve the MINLP problem based on a combination of the echo state network (ESN) learning and reinforcement learning techniques. The echo state learning is shown to be able to predict the optimal movements for the drone hotspots with nearly-constant, low computational complexity in dynamic network environments.
- We develop a new event-driven, universal broadband simulator called SimBAG for integrated aerial-ground wireless networking. An extensive simulation campaign has been conducted based on SimBAG, which proves the great potential of THz-band communications from the network's point of view. Results indicate that significantly (4–6 times) *higher* throughput can be achieved by THz-

band communications in mobile than in static networks. The remainder of the paper is organized as follows. In Section II, we discuss the related works. In Section III, we present the system model and problem formulation. The distributed algorithm design is described in Section IV, and in Section V we discuss the development of SimBAG and analyze the performance evaluation results. Finally, we draw the main conclusions in Section VI.

II. RELATED WORK

MmWave and THz band communications have been explored for wide variety of applications over the last decade [8], [12], [35]–[42]. For example, in [35] the authors propose an algorithm of QoS-aware bandwidth allocation and concurrent scheduling to achieve higher network throughput in THz wireless backhaul networks. In [8], Chaccour et al. explore the potential of THz in VR applications to provide high-data-rate low-latency communications. In [36], the authors study the feasibility of wireless communications in the terahertz bands for four practical altitude-dependant applications. In [37] Barati et al. study the discovery latency and energy consumption caused by beam discovery in mmWave-band communications. In [38] the authors propose a novel risk-based framework for rate optimization and reliable performance for THz-enabled wireless VR networks. The readers are referred to [12], [39]–[42] and references therein for a good survey a main results in this area. None of these works have studied the effects of node mobility on THz-band communications in drone networks. Moreover, while these works have been focusing on either mmWave or THz frequency band, we study joint spectrum access and flight control in drone networks across the microwave, mmWave and THz bands.

Drone-assisted spectrum access has drawn significant research attention [26], [43]–[56]. For example, in [43] Qiu et al. propose an SDN-enabled hierarchical network architecture by integrating the low- and high-altitude platform with the terrestrial cellular networks. In [44] the authors propose a privacy preserving secure spectrum trading and sharing based on blockchain technology to resolve the security issues in wireless UAV networks. In [45] the authors propose a tractable three-dimensional (3D) spatial model for evaluating the average downlink performance of unmanned aerial vehicle (UAV) networks in the mmWave bands. Zhu et al. realize in [46] flexible coverage by exploring 3D beamforming for mmWave UAV communications with a phased uniform planar array. In [47], Gapeyenko et al. investigate the use of UAVs to mitigate the impact of blockage on the backhaul links. In [48], the authors propose a fast beam tracking scheme to achieve high-quality communications in the mmWave band. Xiao et al. explore in [49] the use of mmWave spatial-division multiple access to improve the cellular network capacity. In [50], the authors evaluate the performance of UAV-assisted mmWave networks in urban environments utilizing access points carried by UAVs. In [51], Feng et al. propose a spectrum management architecture and evaluate the performance of the proposed mmWave based wireless backhaul in UAV-assisted cellular networks. Significant efforts have also been made in existing

literature to enable UAV applications in the THz frequency [26], [52]–[55]. In [26] Xia et al. propose to use multi-array antennas to achieve higher data rate with extended communication range for wireless links in the sub-THz bands; in [52] intelligent reflective surface (IRS) is adopted to enhance the data rate for THz-band UAV communications; in our prior work [54], we propose a mobility-resilient beamforming scheme for communications between flying UAVs by dynamically controlling the beamwidth for the mmWave/THz links. The readers are referred to [56] for an excellent survey of the main results in this area.

Reinforcement learning based UAV control has recently received significant attention in the scientific literature [54], [55], [57]–[66]. In [57] Hu et al. propose a distributed sense-and-send protocol to coordinate UAVs for sensing and transmission and use reinforcement learning for trajectory control and resource management. In [58] the authors propose a reinforcement learning scheme to maximize the uplink sum rate in the absence of vital information such as user locations and transmission power. The authors of [59] study the optimal deployment of UAVs based on Q-learning to maximize the real-time downlink capacity. In [60] Zhou et al. propose a deep reinforcement learning (DRL) based dynamic channel allocation scheme by integrating DRL with long short-term memory (LSTM) to learn from the past experience. Similarly, in [61] the authors propose a DRL based method to maximize the energy efficiency of UAV systems by jointly considering the communication coverage, fairness and connectivity. A novel framework for dynamic UAV deployment based on Q-learning is proposed in [62]. In [63] Tafinsev et al. study dynamic associations based on RL for UAV applications in cellular networks with integrated access and backhaul (IAB). In our prior works [54] [55], we proposed an ESN based solution to reduce the link outage probability caused by the movement of drones considering different levels of mobility uncertainties. The readers are referred to [64]–[66] and references therein for an excellent survey of the main results in this field.

Different from the above works, in this paper we focus on a new network scenario called FlyTera, where drone hotspots and ground users are allowed to operate in the microwave, mmWave and THz bands, and study how different spectrum bands complement each other in mobile blockage-rich environments. We use ESN learning to predict the achievable rate of the user by jointly considering the drone locations and design RL-based ESN to optimize their flight trajectories.

III. SYSTEM MODEL AND PROBLEM FORMULATION

In FlyTera there are a set of ground and drone base stations collaboratively providing access services to ground users in the microwave band f^{mc} , mmWave band f^{mm} and THz band f^{tz} , as shown in Fig. 1. Define the set of spectrum bands \mathcal{F} as $\mathcal{F} \triangleq \{f^{mc}, f^{mm}, f^{tz}\}$. Denote \mathcal{B}_{grd} , \mathcal{B}_{fly} and \mathcal{U} as the sets of the ground base stations (GBS), flying drone base stations (FBS) and users, respectively. Let \mathcal{B} represent the set of all the base stations and $\tilde{\mathcal{B}}$ the set of all the nodes, i.e., $\mathcal{B} = \mathcal{B}_{\text{grd}} \cup \mathcal{B}_{\text{fly}}$ and $\tilde{\mathcal{B}} = \mathcal{B} \cup \mathcal{U}$. Our objective is to, given the blockage distribution in the network, maximize the network spectral efficiency by

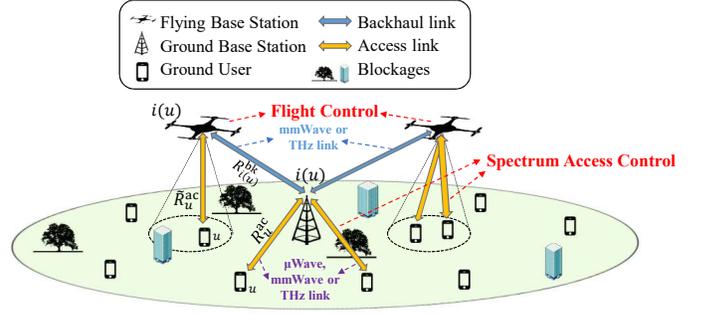


Fig. 1: FlyTera System Model.

jointly controlling the flight of the FBSs, their association with the ground users as well as the spectrum bands used by the users. Next we describe the blockage, link, spectrum access models and mobility models sequentially. The key notations are summarized in Table I for the reader’s convenience.

A. Blockage Model

Let \mathcal{K} represent the set of blockages in the network. As illustrated in Fig. 2, each blockage $k \in \mathcal{K}$ is represented as a rectangle of dimensions $L_k \times W_k \times H_k$, with L_k , W_k and H_k being the length, width and height of the blockage, respectively. Denote C_k as the center of blockage k . The orientation of blockage k , denoted as θ_k , is considered to be uniformly distributed in $[0, 2\pi]$. Define $\mathbf{P}_k^{\text{blk}}(C_k, L_k, W_k, H_k, \theta_k)$ as the set of points contained in blockage k .

Let $\text{cod}_i = (x_i, y_i, z_i)$ denote the coordinate vector of node $i \in \tilde{\mathcal{B}}$ (the phone in Fig. 2), with x_i , y_i and z_i being the x-, y- and z-axis components, respectively. Similarly, denote cod_j as the coordinate vector of node $j \in \tilde{\mathcal{B}}$ (the drone in Fig. 2). Further define $\mathbf{P}_{ij}^{\text{seg}}$ as the point set of the segment connecting nodes i and j . Finally, use $\text{I}(\text{cod}_i, \text{cod}_j, k)$ to indicate whether blockage k is blocking the link between nodes i and j , with $\text{I}(\text{cod}_i, \text{cod}_j, k) = 1$ if yes, i.e., $\mathbf{P}_{ij}^{\text{seg}} \cap \mathbf{P}_k^{\text{blk}}(C_k, L_k, W_k, H_k, \theta_k) \neq \phi$, and $\text{I}(\text{cod}_i, \text{cod}_j, k) = 0$ otherwise.

Then, given the set \mathcal{K} of blockages, the total number of blockages in the link between nodes i and j , denoted as $K^{i,j}$, can be expressed as

$$K^{i,j} = \sum_{k \in \mathcal{K}} \text{I}(\text{cod}_i, \text{cod}_j, k), \forall i, j \in \tilde{\mathcal{B}}. \quad (1)$$

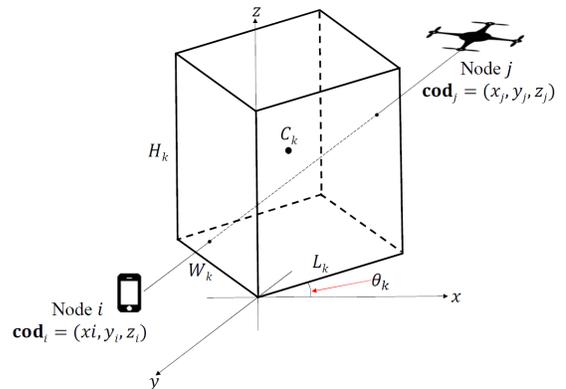


Fig. 2: Illustration of the blockage model in FlyTera.

Notation	Physical Meaning
f_{mc}	Frequency of microwave frequency band
f_{mm}	Frequency of millimeter-wave frequency band
f_{tz}	Frequency of terahertz frequency band
\mathcal{F}	Set of spectrum bands
\mathcal{B}_{grd}	Set of ground base stations
\mathcal{B}_{fly}	Set of flying drone base stations
\mathcal{U}	Set of users
$\tilde{\mathcal{B}}$	Set of all base stations
$\tilde{\mathcal{B}}$	Set of all nodes
\mathcal{K}	Set of blockages
L_k, W_k, H_k	Length, width, height of a blockage $k \in \mathcal{K}$
C_k, θ_k	Center and orientation of a blockage $k \in \mathcal{K}$
$\mathbf{P}_k^{blk}(\cdot)$	Set of points contained in blockage k
\mathbf{cod}_i	x-, y-, z- coordinate vector of node $i \in \tilde{\mathcal{B}}$
\mathbf{cod}_j	x-, y-, z- coordinate vector of node $j \in \tilde{\mathcal{B}}$
\mathbf{P}_{ij}^{seg}	Point set of segment connecting nodes i and j
$I(\cdot)$	Indicator function
$K^{i,j}$	Total number of blockages in the link i, j
$L_{i,j}$	Path loss between nodes i and j
β_0	Per-blockage absorption coefficient
$\alpha_{i,j}(f)$	Path loss exponent for link i, j in frequency band f
C	Speed of light
$d_{i,j}$	Distance between nodes i and j
γ_u^{mc}	Microwave band link SINR for user u
γ_u^{mm}	Millimeter-wave band link SINR for user u
γ_u^{tz}	Terahertz band link SINR for user u
$i(u)$	Serving base station of user u
$P_{i(u)}$	Transmission power of serving base station $i(u)$
P_j	Transmission power of interfering base station $j \neq i(u)$
$\mathcal{U}_{i(u)}$	Set of users served by serving base station $i(u)$
N_0	Noise power
$R_u(\cdot)$	Rate achievable by user u
G	Transmit gain of base station
\tilde{G}	Receive gain of user
θ	Offset angle (boresight) of user u 's antenna
θ'	Offset angle (boresight) of drone base station antenna
$\psi(\cdot, \cdot)$	Frequency selection function
R^{ac}, R^{bk}	Access and backhaul link rates
ζ	Association vector
\mathcal{T}	Total simulation duration

TABLE I: Summary of Key Notations.

B. Link Model

In this section we describe the interference model focusing on ground users in \mathcal{U} , while the model can be derived similarly for flying base stations in \mathcal{B}_{fly} . To this end, we first describe the path loss model.

Path Loss. Denote $L_{i,j}(f)$ as the path loss between nodes $i, j \in \tilde{\mathcal{B}}$ operating in frequency band $f \in \mathcal{F}$. Then $L_{i,j}(f)$ can be modelled as in [67]:

$$L_{i,j}(f) = \beta_0^{K_{i,j}} \left(\frac{4\pi f}{C} \right)^2 (d_{i,j})^{\alpha_{i,j}(f)} \quad (2)$$

where C is the speed of light, $\alpha_{i,j}(f)$ is the path-loss exponent for the link between nodes i and j in frequency band f , $K_{i,j}$ defined in (1) represents the number of blockages along the link, $\beta_0 \in [0, 1]$ is the per-blockage absorption coefficient [9] [68], and finally $d_{i,j} = d_{i,j}(\mathbf{cod}_i, \mathbf{cod}_j)$ denotes the distance between nodes i and j , i.e.,

$$d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, \quad (3)$$

given coordinate vectors \mathbf{cod}_i and \mathbf{cod}_j defined in Section III-A for nodes i and j , respectively.

Microwave-Band Link. Denote γ_u^{mc} as the SINR of ground user $u \in \mathcal{U}$ if it receives on the microwave frequency band f^{mc} , then γ_u^{mc} can be expressed as

$$\gamma_u^{mc} = \frac{P_{i(u)}^{mc} L_{i(u),u}(f^{mc})}{\sum_{j \in \mathcal{B}^{mc}/i(u)} P_j^{mc} L_{j,u}(f^{mc}) + N_0^{mc}} \quad (4)$$

where $\mathcal{B}^{mc} \subset \mathcal{B}$ represents the set of BSs operating on this band, and $i(u) \in \mathcal{B}^{mc}$ represents the serving BS of user $u \in \mathcal{U}$; $P_{i(u)}^{mc}$ and P_j^{mc} are the transmission power of the serving BS $i(u)$ and interfering BS $j \in \mathcal{B}^{mc}/i(u)$, respectively; $L_{i(u),u}(f^{mc})$ and $L_{j,u}(f^{mc})$ are the path loss defined in (2), and finally N_0^{mc} is the power of noise at node u in the microwave band and is expressed as $N_0^{mc} = N_{\text{thermal}} = k_B T B^{mc}$, where k_B is the Boltzmann constant, T is the temperature and B^{mc} is the bandwidth of the microwave band. Let $\mathcal{U}_{i(u)}^{mc} \in \mathcal{U}$ represent the set of users served by BS $i(u)$ in this band and consider that its transmission time is shared among its users based on time-division multiple access (TDMA). Then the rate achievable by user u in this band, denoted as $R_u(f^{mc})$, can be written as

$$R_u(f^{mc}) = \frac{1}{|\mathcal{U}_{i(u)}^{mc}|} B^{mc} \log_2(1 + \gamma_u^{mc}), \quad (5)$$

where $|\cdot|$ represents the cardinality of a set, i.e., the number of users served by BS $i(u)$ in this frequency band for our case; coefficient $\frac{1}{|\mathcal{U}_{i(u)}^{mc}|}$ denotes the fraction of the time allocated to user $u \in \mathcal{U}_{i(u)}^{mc}$.

MmWave-Band Link. Let $P_{i(u)}^{mm}$ denote the transmission power of the serving base station of user u (i.e., $i(u)$) in the mmWave band. Further denote $\mathcal{U}_{i(u)}^{mm} \in \mathcal{U}$ as the set of users served by BS $i(u)$ and $|\mathcal{U}_{i(u)}^{mm}|$ as the number of users in $\mathcal{U}_{i(u)}^{mm}$. Different from the microwave band, where the BS serves its users based on TDMA, in mmWave band the BS is able to serve the users simultaneously with the directional mmWave-band links. Let $P_{i(u),u'}^{mm}$ represent the transmission power of BS $i(u)$ allocated to user $u' \in \mathcal{U}_{i(u)}^{mm}$, then we have

$$\sum_{v \in \mathcal{U}_{i(u)}^{mm}} P_{i(u),v}^{mm} \leq P_{i(u)}^{mm}, \quad \forall u \in \mathcal{U}^{mm}. \quad (6)$$

The received SINR of user $u \in \mathcal{U}$ in this band, denoted as γ_u^{mm} , can then be written as

$$\gamma_u^{mm} = \frac{P_{i(u),u}^{mm} L_{i(u),u}(f^{mm}) G_{\max}^{mm} \tilde{G}_{\max}^{mm}}{\sum_{u' \in \mathcal{U}^{mm}/u} P_{i(u'),u'}^{mm} L_{i(u'),u}(f^{mm}) G_{i(u'),u}^{mm} \tilde{G}_{u,i(u')}^{mm} + N_0^{mm}} \quad (7)$$

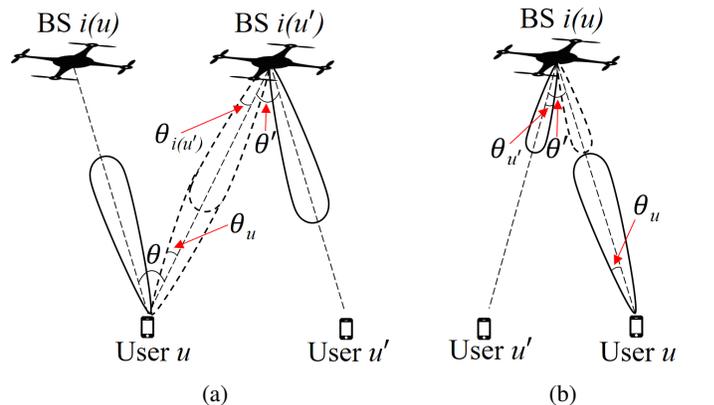


Fig. 3: Interference model for the mmWave-band links: (a) inter-BS interference; (b) intra-BS interference.

where \mathcal{U}^{mm} represents the set of all the users operating in the mmWave band, and N_0^{mm} is the power of noise in this band at each user. In (7), $L_{i(u),u}(f^{\text{mm}})$ represents the path loss between BS $i(u)$ and user u in the mmWave band; $G_{i(u'),u}^{\text{mm}}$ and $\tilde{G}_{u,i(u')}^{\text{mm}}$ represent the transmit gain of BS $i(u')$ and receive gain of user u respectively; $G_{\text{max}}^{\text{mm}}$ and $\tilde{G}_{\text{max}}^{\text{mm}}$ denote the maximum transmit gain of BSs and maximum receive gain of users, respectively. Denote the resulting link rate as $R_u(f^{\text{mm}})$ for user u , then we have

$$R_u(f^{\text{mm}}) = B^{\text{mm}} \log_2(1 + \gamma_u^{\text{mm}}), \quad (8)$$

where B^{mm} is the bandwidth of the mmWave frequency band.

We consider sectorized interference model as in [69] to determine the transmit and receive gains, i.e., $G_{i(u'),u}^{\text{mm}}$ and $\tilde{G}_{u,i(u')}^{\text{mm}}$ in (7). As illustrated in Fig. 3, let $\theta \in [-\pi, \pi]$ denote the offset angle of the antenna boresight direction of user u with respect to the reference direction, and θ' as the offset angle for the drone base station. Here, the reference direction refers to the direction along which the transmitting and receiving beams are perfectly aligned, as indicated by the dashed beams in Fig. 3. Denote θ_u and $\theta_{i(u')}$ as the beamwidth of user u and BS $i(u')$, respectively. Then, the transmit gain for BS $i(u')$ and receive gain of user u can be determined as follows, taking Fig. 3(a) as an example:

$$G_{i(u'),u}^{\text{mm}} = \begin{cases} G_{\text{max}}^{\text{mm}}, & \text{if } \theta' \leq \theta_{i(u')} \\ G_{\text{min}}^{\text{mm}}, & \text{otherwise} \end{cases} \quad (9)$$

for the transmit gain, and

$$\tilde{G}_{u,i(u')}^{\text{mm}} = \begin{cases} G_{\text{max}}^{\text{mm}}, & \text{if } \theta \leq \theta_u \\ G_{\text{min}}^{\text{mm}}, & \text{otherwise} \end{cases} \quad (10)$$

for the receive gain.

THz-Band Link. The SINR achievable by a user in the THz band can be derived similarly as in (7)-(10) for the mmWave band, except that only LOS transmissions will be considered because of the significantly higher path loss. Then, we have $K_{i,j} = 0$ in (2) and the overall path loss can be rewritten as

$$L_{i,j}(f^{\text{tz}}) = \left(\frac{4\pi f^{\text{tz}}}{C} \right)^2 (d_{i,j})^{\alpha_{i,j}(f^{\text{tz}})} e^{A_{\text{abs}}(f^{\text{tz}}, d_{i,j})}, \quad (11)$$

where $d_{i,j}$ defined in (3) is the distance between nodes i and j , and $\alpha_{i,j}(f^{\text{tz}})$ represents the path loss exponent in the THz band and $A_{\text{abs}}(f^{\text{tz}}, d_{i,j})$ is molecular absorption factor at frequency f^{tz} and distance $d_{i,j}$. Finally, denote the resulting SINR as γ_u^{tz} for user u and the corresponding link rate as $R_u(f^{\text{tz}})$.

C. Spectrum Access Model

Consider single-band spectrum access for the ground users and multi-band spectrum access for the base stations, i.e., each user is allowed to use at most one frequency band at the same time, while each base station is able to serve multiple users in different frequency bands. Then we have

$$\sum_{f \in \mathcal{F}} \psi(u, f) \leq 1, \quad \forall u \in \mathcal{U} \quad (12)$$

where $\psi(u, f)$ is the frequency selection function, with $\psi(u, f) = 1$ if frequency band f is used by user u , and $\psi(u, f) = 0$ otherwise. Then the overall access link rate of user $u \in \mathcal{U}$, denoted as R_u^{ac} , can be expressed as

$$R_u^{\text{ac}} = \sum_{f \in \mathcal{F}} \psi(u, f) R_u(f), \quad (13)$$

where $R_u(f)$ is defined in Section III-B, with $f \in \mathcal{F}$. Since the aggregate access link rate of each FBS should not exceed the rate of the backhaul link, i.e., the link between FBS and the ground network infrastructure, the adjusted access link rate, denoted as \tilde{R}_u^{ac} for user $u \in \mathcal{U}$, can be given as

$$\tilde{R}_u^{\text{ac}} = \min \left(\underbrace{\sum_{v \in \mathcal{U}_{i(u)}} R_v^{\text{ac}}, R_{i(u)}^{\text{bk}}}_{\text{Minimum of access and backhaul link rates}}, \underbrace{\frac{R_u^{\text{ac}}}{\sum_{v \in \mathcal{U}_{i(u)}} R_v^{\text{ac}}}}_{\text{Proportional rate allocation among users}} \right) \quad (14)$$

where $R_{i(u)}^{\text{bk}}$ is the backhaul link rate of user u 's serving base station, i.e., $i(u)$ and $\mathcal{U}_{i(u)}$ is the set of users sharing base station $i(u)$. In (14), the first item on the right-hand side is used to determine the minimum rate of the access and backhaul links, and the objective of the second item is to allocate the resulting minimum rate among the users sharing the same backhaul link.

D. Problem Statement

Finally, the control objective of FlyTera is to maximize the aggregate rate of all the users in \mathcal{U} , by jointly controlling the flight of the flying base stations in \mathcal{B}_{fly} , their association with the ground users as well as the spectrum band selection of the users, under the constraints of single-band access for the users and wireless backhaul links for the flying base stations. Let $\mathbf{cod} = (\mathbf{cod}_i)_{i \in \tilde{\mathcal{B}}}$ represent the coordinate vector of all the nodes in the network, and $\boldsymbol{\psi} = (\psi(u, f))_{u \in \mathcal{U}}^{f \in \mathcal{F}}$ denote the spectrum band selection vector of the users. Further denote $\boldsymbol{\zeta} = (\zeta_{ui})_{u \in \mathcal{U}}^{i \in \mathcal{B}}$ as the association vector, with $\zeta_{ui} = 1$ if user u is associated with BS i and $\zeta_{ui} = 0$ otherwise. If we consider single-home association for the users, i.e., each user is allowed to be associated to at most one base station at the same time, then we have

$$\sum_{i \in \mathcal{B}} \zeta_{ui} \leq 1, \quad \forall u \in \mathcal{U}. \quad (15)$$

The FlyTera control problem can then be formulated as

$$\begin{aligned} \text{Given :} & \quad \mathcal{U}, \mathcal{B}_{\text{fly}}, \mathcal{B}_{\text{grd}}, \mathcal{F} \\ \text{Maximize} & \quad \sum_{u \in \mathcal{U}} \tilde{R}_u^{\text{ac}}(\mathbf{cod}, \boldsymbol{\psi}, \boldsymbol{\zeta}) \\ \text{cod}, \boldsymbol{\psi}, \boldsymbol{\zeta} & \\ \text{Subject to :} & \quad (6), (12), (15) \end{aligned} \quad (16)$$

where $\tilde{R}_u^{\text{ac}}(\mathbf{cod}, \boldsymbol{\psi}, \boldsymbol{\zeta}) = \tilde{R}_u^{\text{ac}}$ defined in (14) is the adjusted access link rate of user u .

The network control problem formulated in (16) is a mixed integer nonlinear nonconvex (MINLP) problem, because of the involved mathematical expression of $\tilde{R}_u^{\text{ac}}(\mathbf{cod}, \boldsymbol{\psi}, \boldsymbol{\zeta})$ and the binary association variables and frequency selection variables. Given an arbitrary such problem, it is still an open problem

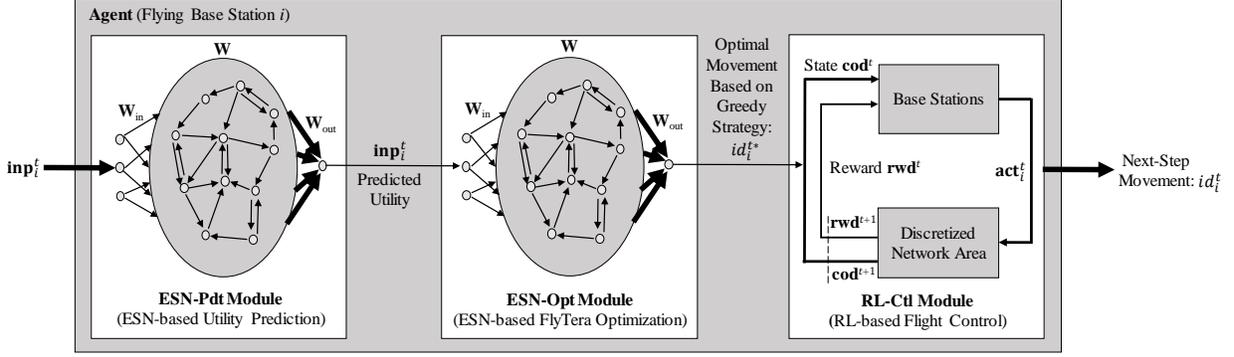


Fig. 4: Diagram of the distributed algorithm design based on a combination of echo state learning and reinforcement learning.

to obtain the globally optimal solution with polynomial computational complexity. Recall in Section I that in this work our objective is to investigate the applicability of THz-band communications in mobile blockage-rich environments and how the THz band can complement the lower-frequency bands. To this end, in next section we solve problem (16) by designing distributed algorithms based on a combination of echo state learning and reinforcement learning, and then evaluate their performance in Section V.

IV. DISTRIBUTED SOLUTION ALGORITHMS

The framework of the distributed solution algorithm design is illustrated in Fig. 4, where there are three major modules, i.e., i) *ESN-based Utility Prediction (ESN-Pdt)*, ii) *ESN-based Utility Optimization (ESN-Opt)*, and iii) *RL-based FlyTera Control (RL-Ctl)*. The objective of *ESN-Pdt* is to predict the utility for each FBS by approximating the mapping from the network control variables, i.e., \mathbf{cod} , ψ and ζ in (16), to the individual utility function $\tilde{R}_u^{\text{ac}}(\mathbf{cod}, \psi, \zeta)$ based on echo state learning. Then, given the predicted utilities each FBS determines its own next-step action based on the *ESN-Opt* module. Finally, the *RL-Ctl* module is used to achieve a tradeoff between exploring and exploiting in favor of higher network spectral efficiency. In this work we design the RL algorithm based on ESN, which is a new reservoir computing technique for recurrent neural networks. Comparing to traditional neural networks, ESN is simpler to train while still able to model the complex time-varying behaviors of dynamical systems [70]–[79]. It is worth pointing out that in this work our focus is on studying how different spectrum bands complement each other in mobile blockage-rich UAV networks, and the designed joint flight and spectrum control framework can also be extended to other learning techniques such as LSTM [80], actor critic methods [81].

A. ESN-Pdt: ESN-based Utility Prediction

In FlyTera with distributed flying drone base stations, it is hard to obtain the complete and up-to-date network status information required for the base stations to derive the exact mathematical expression of $\tilde{R}_u^{\text{ac}}(\mathbf{cod}, \psi, \zeta)$. To address this challenge, in this work we approximate the individual utility function $\tilde{R}_u^{\text{ac}}(\mathbf{cod}, \psi, \zeta)$ based on ESN. Roughly speaking, as shown in Fig. 4, the objective of an ESN is to

model approximately the mapping from the input signals to the output signals of a system, by training its input weights \mathbf{W}_{in} , the reservoir weights \mathbf{W} and output weights \mathbf{W}_{out} using a sigmoidal transfer function (e.g., hyperbolic tangent). Next we first describe the *ESN-Pdt* module (the left block in Fig. 4) design in FlyTera.

The *ESN-Pdt* module consists of four components: *Agent*, *Input*, *Action* and *Reward Function*. In *ESN-Pdt*, the *Agent* refers to individual BSs in \mathcal{B}_{fly} , i.e., each BS is endowed with an *ESN-Pdt* module for approximating its own utility function. Then, in each time slot $t \in \mathcal{T}$, each BS $i \in \mathcal{B}_{\text{fly}}$ feeds an *Input* (denoted as \mathbf{inp}_i^t) and a candidate *Action* (denoted as \mathbf{act}_i^t) to its *ESN-Pdt* module, which will then output the *Reward Function* value of the BS.

Input Design. The input of BS i 's *ESN-Pdt* module in time slot t , defined as $\mathbf{inp}_i^t \triangleq \{\mathbf{cod}_{-i}^t, \psi^t, \zeta^t\}$, comprises the locations of all the other BSs $\mathbf{cod}_{-i}^t = (\mathbf{cod}_j^t)_{j \in \mathcal{B}_{\text{fly}}/i}$ with \mathbf{cod}_j^t being the coordinate vector of BS j in time slot t , the association profile of the ground users ζ^t as well as their spectrum band selection strategies ψ^t (confer Section III for the definitions of \mathbf{cod}_j^t , ψ^t and ζ^t)¹. The dimension of \mathbf{inp}_i^t increases quadratically with the scale of the network because of the association vector ζ^t . This can slow down the training of the *ESN-Pdt* module and hence degrade the utility approximation accuracy in large-scale networks. To address this challenge, we reform \mathbf{inp}_i^t by reducing the number of primal variables in the input vector based on the following three simple but effective policies. First, each ground user $u \in \mathcal{U}$ is associated to its nearest base station, i.e., the serving base station $i(u)$ is selected so that

$$i(u) = \arg \min_{i' \in \mathcal{B}} d_{i'u}(\mathbf{cod}_{i'}, \mathbf{cod}_u), \forall u \in \mathcal{U}, \quad (17)$$

where $d_{i'u}(\cdot, \cdot)$ calculates the distance between two nodes as defined in (3). Second, two distance thresholds are adopted, denoted as d_0^{mc} and d_0^{tz} with $d_0^{\text{tz}} < d_0^{\text{mc}}$, for the microwave and THz bands, respectively. Then, the microwave band will be selected for user u if $d_{u,i(u)} < d_0^{\text{mc}}$, i.e., the distance between user u and its serving base station $i(u)$ determined based on the first policy, satisfies $d_{u,i(u)} \geq d_0^{\text{mc}}$; if $d_{u,i(u)} \leq d_0^{\text{tz}}$, the THz band will be selected; otherwise, user u will use the mmWave band if $d_0^{\text{tz}} < d_{u,i(u)} < d_0^{\text{mc}}$. Finally, the

¹In previous sections, superscript t has been omitted for notational convenience.

network area is divided into a number $N_x \times N_y \times N_z$ of three-dimensional rectangles, each with $\frac{L_x}{N_x}$, $\frac{L_y}{N_y}$ and $\frac{L_z}{N_z}$ for width, length and height, respectively. Denote \mathcal{N} as the set of the resulting rectangles. Each rectangle $n \in \mathcal{N}$ is represented using a vector $\text{rect}_n = (\text{cod}_n, id_n)$, where cod_n is the coordinate vector of the center point of the rectangle, and $id_n = 0, 1, \dots, N_x \times N_y \times N_z - 1$ is the rectangle index.

Based on the first two policies, the association and spectrum band selection vectors (i.e., ψ^t and ζ^t) can be determined given the coordinates of all the BSs (i.e., cod^t) in each time slot t . As a result, cod_{-i}^t is the only primal variable in inp_i^t . The dimension of inp_i^t can be further reduced based on Policy 3 by mapping each component of cod_{-i}^t to the index of the corresponding rectangle. Finally, the input of the *ESN-Pdt* module can be rewritten as $\text{inp}_i^t = (id_j^t)_{j \in \mathcal{B}_{\text{fly}}/i}$, with id_j^t being the index of BS j 's rectangle in time slot t .

Action and Reward. Given input inp_i^t for the *ESN-Pdt* module for BS i in time slot t , BS i makes its action decisions and observes an output of the action. To this end, BS i chooses to move to a new rectangle in \mathcal{N} except those occupied by other BSs. Then, the set of actions for BS i , denoted as act_i^t for time slot t , can be written as

$$\text{act}_i^t = \{id_m | m \in \mathcal{N} \setminus \{n^t(j), j \in \mathcal{B}_{\text{fly}}/i\}\}, \quad (18)$$

where $n^t(j)$ represents the rectangle index of BS j in time slot t . The corresponding reward, denoted as $\text{rwd}_i^t(id_m)$, is determined by the aggregate rate achievable by the users it serves at new location id_m , i.e.,

$$\text{rwd}_i^t(id_m) = \sum_{v \in \mathcal{U}_i(id_m)} \tilde{R}_v^{\text{ac}}, \quad (19)$$

where \tilde{R}_v^{ac} defined in (14) is the adjusted access rate of user v and $\mathcal{U}_i(id_m)$ is the set of users served by BS i at rectangle id_m .

ESN-Pdt Training. During the *training phase*, we move each FBS to different coordinates and measure the resulting utility i.e., the sum access rate defined in (19) and use it as the target rate denoted as $\text{rwd}_{\text{tgt}}^t$ for time slot t . The objective of ESN training is to learn a model that can approximate the mapping between the input data and target rate $\text{rwd}_{\text{tgt}}^t$. In this work, this is accomplished by setting the cost function of the *ESN-Pdt* module as the root-mean-square error (RMSE) between rwd_i^t and $\text{rwd}_{\text{tgt}}^t$ as follows, for each FBS $i \in \mathcal{B}_{\text{fly}}$,

$$E(\text{rwd}_i^t, \text{rwd}_{\text{tgt}}^t) = \frac{1}{N_{\text{out}}} \sum_{n=1}^{N_{\text{out}}} \sqrt{\frac{1}{|\mathcal{T}|} \sum_{t=1}^{|\mathcal{T}|} (\text{rwd}_i^t - \text{rwd}_{\text{tgt}}^t)^2}, \quad (20)$$

where N_{out} is the dimension of the output units of the ESN and $|\cdot|$ represents the cardinality of a set, and rwd_i^t is the predicted reward for FBS i in time slot t . The collection of the training data and the validation of the trained ESN model will be discussed in Section V.

B. ESN-Opt: ESN-based FlyTera Optimization

The job of the *ESN-Opt* module (the middle block in Fig. 4) is to determine the optimal next-step location for each BS,

Algorithm 1: FlyTera Algorithm

Data: FBS Coordinates $\text{cod}_i = (x_i, y_i, z_i); \forall i \in \mathcal{B}_{\text{fly}}$, Total duration of simulation \mathcal{T}

Result: Action act_i^{t+1} for time slot $t+1 \in \mathcal{T}$ determined by *RL-Ctl* module based on the combination of *ESN-Opt* and *ESN-Pdt* modules

- 1 **Initialization:** Set of actions act_i defined in (18)
- 2 **while** $t \in \mathcal{T}$ **do**
- 3 **ESN-Pdt: ESN-based Utility Prediction**
- 4 **for** each FBS $i \in \mathcal{B}_{\text{fly}}$ **do**
- 5 **for** each rectangle index $id_m, m \in \mathcal{N}$ **do**
- 6 Calculate the reward based on (19) and use it as target rate $\text{rwd}_{\text{tgt}}^t$;
- 7 Train an ESN model (i.e., *ESN-Pdt*) that can minimize RMSE between predicted rate rwd_i^t and target rate $\text{rwd}_{\text{tgt}}^t$ as defined by (20);
- 8 **end**
- 9 **end**
- 10 **ESN-Opt: ESN-based FlyTera Optimization**
- 11 **for** each FBS $i \in \mathcal{B}_{\text{fly}}$ **do**
- 12 Predict reward for each candidate based on *ESN-Pdt* module as defined in (21);
- 13 Train an ESN model (i.e., *ESN-Opt*) that can minimize RMSE between predicted location index id^t and target location index id_{tgt}^t .
- 14 **end**
- 15 **RL-Ctl: Reinforcement Learning Based Flight Control**
- 16 **for** each FBS $i \in \mathcal{B}_{\text{fly}}$ **do**
- 17 Determine the next-step best location based on the combination of *ESN-Pdt* and *ESN-Opt* modules;
- 18 Determine act_i^{t+1} based on (22).
- 19 **end**
- 20 **end**

given the locations of all the other BSs. The agents and inputs of this module is the same as in *ESN-Pdt*, that is, *ESN-Opt* is also operated in individual flying BSs, and each BS takes the location information of the other BSs (i.e., inp_i^t defined above) as its input. Differently, the action set of BS i , denoted $\widetilde{\text{act}}_i^t$ contains only single rectangle in each time slot t (its current rectangle), i.e., $\widetilde{\text{act}}_i^t = \{id_i^t\}$. The reward, denoted as $\widetilde{\text{rwd}}_i^t$, is the maximum utility that BS i may achieve by moving to a new rectangle in next time slot, i.e.,

$$\widetilde{\text{rwd}}_i^t = \max_{id \in \widetilde{\text{act}}_i^t} \text{rwd}_i^t(id) \quad (21)$$

where rwd_i^t and act_i^t are the reward and the action set of BS i 's *ESN-Pdt* module discussed in Section IV-A. Denote the resulting optimal next-step rectangle for BS i as id_i^{t*} .

C. RL-Ctl: Reinforcement Learning Based Flight Control

Based on a combination of the *ESN-Pdt* and *ESN-Opt* modules discussed above, each FBS determines its own best location for the next time slot. This may lead to a local optimum of the FlyTera control problem (16), which is not desirable. In favor of high network spectral efficiency, in this work we use reinforcement learning to guide the *exploration* and *exploitation* in the flight control. Reinforcement learning (RL) [82] has been widely used to solve very complex problems that cannot be solved by conventional techniques. The

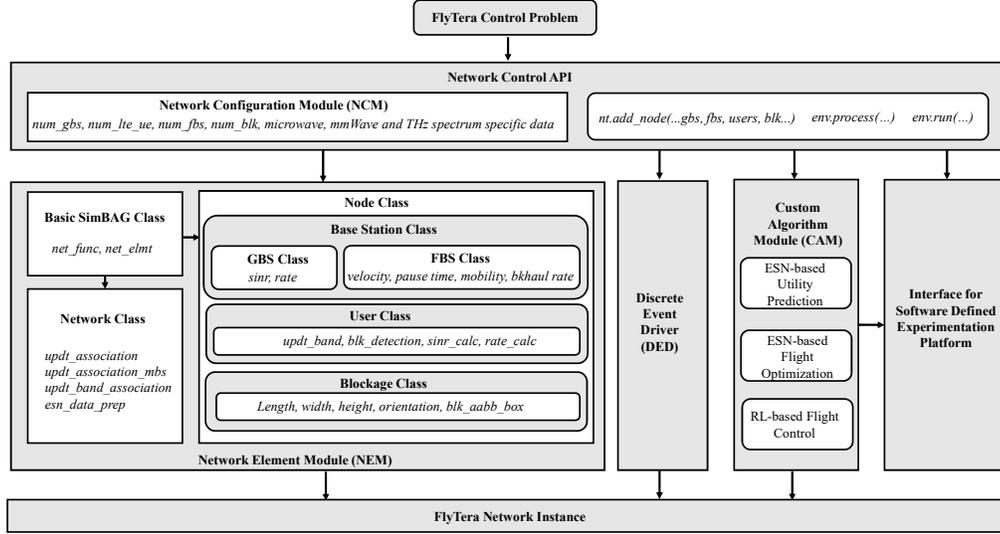


Fig. 5: Architecture of SimBAG: a simulator for broadband aerial-ground wireless networks.

same as above discussed ESN modules, as shown in Fig. 4 (the right block), the RL agents are also the flying base stations, the environment is the discretized network area. The state of each agent i is denoted as cod_i^t in time slot t , and the feedback reward from the environment is denoted as rwd_i^t . Then for any action act_i^t taken by the agent, let cod_i^{t+1} and rwd_i^{t+1} denote the corresponding state and reward at time $t+1$. The output of the *RL-Ctl* module is the next optimal action denoted as act_i^{t*} for the agent i in time slot t . In this paper, we design the *RL-Ctl* module based on ϵ -greedy exploration strategy [82]. The pseudocode of FlyTera Algorithm is summarized in Algorithm 1.

Theorem 1. *If mixed strategies are adopted by the FBSs, the distributed algorithm proposed in this section converges to a stationary network operating point at which no FBS has incentive to fly to other locations if the other FBSs do not.*

Proof. The theorem can be proved based on the convergence framework developed in [79]. Specifically, given the finite set of actions act_i^t defined in (18), let $\Delta(\text{act}_i^t)$ represent the set of all probability distributions over the elements of act_i^t , and $\pi_i = [\pi_i(id_1), \dots, \pi_i(id_{|\mathcal{N}|})]$ with $\pi_i \in \Delta(\text{act}_i^t)$ denoting the probability distribution used by FBS $i \in \mathcal{B}_{\text{fly}}$ to select an action from its action set act_i^t , and $|\mathcal{N}|$ being the cardinality of the set of rectangles defined in Section IV-A. Then, the mixed strategy profile for FBS i , denoted as $\pi_i^* \in \Delta(\text{act}_i^t), \forall i \in \mathcal{B}_{\text{fly}}$, can be given as $\pi_i^* = [\pi_i^*(id_1), \dots, \pi_i^*(id_{|\mathcal{N}|})]$. The flight control and spectrum access problem (16) can then be reformulated as a non-cooperative game, and to prove this theorem it is sufficient to show that the game converges to a mixed Nash Equilibrium (MNE) with mixed strategy probability [79], [83]. To this end, we need to show the following condition holds for a mixed strategy profile $\pi_i^* = [\pi_i^*(id_1), \dots, \pi_i^*(id_{|\mathcal{N}|})] = (\pi_i^*, \pi_{-i}^*)$ [79], i.e., $\tilde{u}_i(\pi_i^*, \pi_{-i}^*) \geq \tilde{u}_i(\pi_i, \pi_{-i}^*)$.

Recall that ϵ -greedy exploration strategy is adopted in FlyTera to ensure that the probability of choosing an action act_i^t is always greater than 0. Then, the probability of FBS

$i \in \mathcal{B}_{\text{fly}}$ choosing an action act_i^t can be given as

$$P(\text{act}_i^t) = \begin{cases} 1 - \epsilon + \frac{\epsilon}{|\text{act}_i^t|}; & \arg \max_{\text{act}_i^t} \text{rwd}(\text{act}_i^t) \\ \frac{\epsilon}{|\text{act}_i^t|}; & \text{a random action} \end{cases} \quad (22)$$

Let $\widetilde{\text{act}}_i^{t+1,*}$ denote the action that results in optimal reward given the optimal mixed strategy (π_i^*, π_{-i}^*) . Then the utility function of FBS i can be given as

$$\begin{aligned} & \tilde{u}_i(\pi_i^*, \pi_{-i}^*) - \tilde{u}_i(\pi_i, \pi_{-i}^*) \\ &= \sum_{\widetilde{\text{act}}_i^{t+1} \in \text{act}_i^t} [\pi_{i, \widetilde{\text{act}}_i^{t+1}}^* \text{U} - \pi_{i, \widetilde{\text{act}}_i^{t+1}} \text{U}] \\ &= \sum_{\widetilde{\text{act}}_i^{t+1} \in \text{act}_i^t} \mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1})] \{ \pi_{i, \widetilde{\text{act}}_i^{t+1}}^* - \pi_{i, \widetilde{\text{act}}_i^{t+1}} \} \\ &= (1 - \epsilon) \left(\mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1,*})] - \mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1})] \right) \end{aligned} \quad (23)$$

where $\text{U} = \sum_{\widetilde{\text{act}}_i^{t+1} \in \text{act}_i^t} u_i(\text{act}_i^{t+1}, \text{act}_{-i}^{t+1}) \pi_{i, \text{act}_i^{t+1}}^*$. (24) is obtained from (22). We can then conclude that $\mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1,*})] - \mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1})] \geq 0$ based on the fact that in FlyTera the optimal action $\text{act}_i^{t+1,*}$ results in optimal $\mathbb{E}[u_i(\text{act}_i^{t+1,*})]$ and the value of $\mathbb{E}[u_i(\widetilde{\text{act}}_i^{t+1})]$ cannot exceed the optimal value and therefore the difference is always greater than or equal to 0. This completes the proof. ■

V. PERFORMANCE EVALUATION

In this section, we first verify the effectiveness of the distributed solution algorithms proposed in Section IV and then analyze the performance of FlyTera. As of today, there are still no publicly available testbed or simulator that can support experiments of integrated aerial-ground wireless networking in the microwave, mmWave and THz bands. In this work we

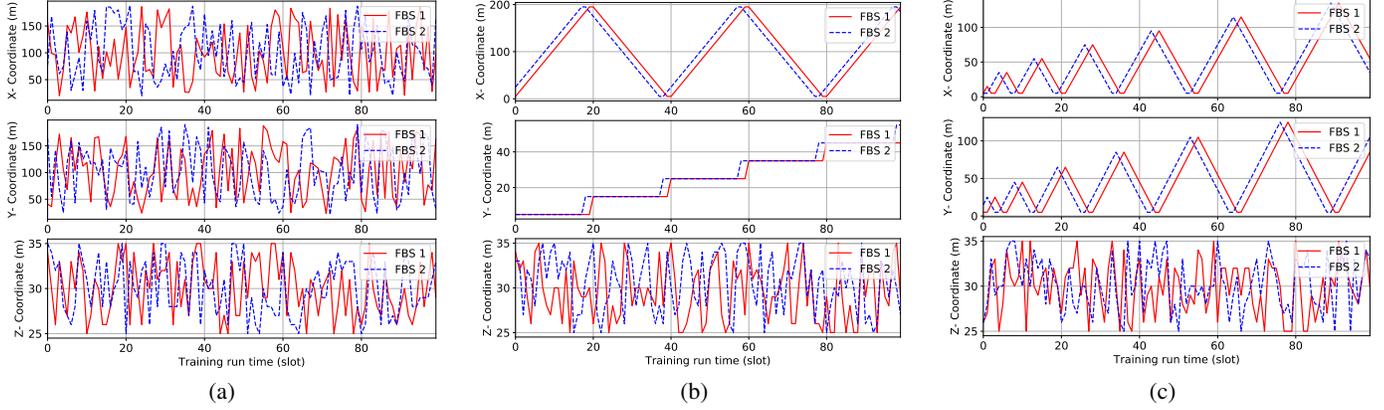


Fig. 7: Examples trajectory of FBS during the training process with (a) random movement, (b) horizontal scanning and (c) diagonal scanning mobility models, respectively.

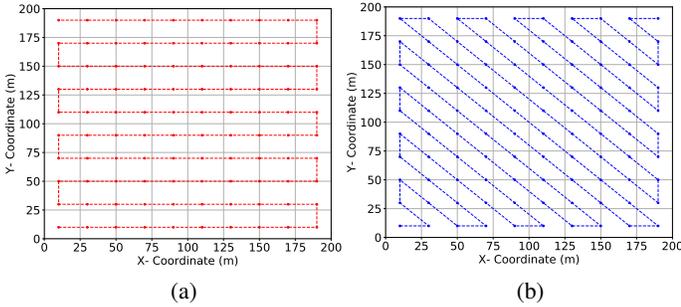


Fig. 6: Illustration of (a) horizontal scanning and (b) diagonal scanning.

conduct simulations over a newly designed simulator called *SimBAG*².

A. SimBAG Design

SimBAG is a Python-based event-driven simulator for broadband integrated aerial-ground wireless networks. As illustrated in Fig. 5, SimBAG comprises four modules: *Network Configuration Module (NCM)*, *Network Element Module (NEM)*, *Discrete Event Driver (DED)*, and *Custom Algorithm Module (CAM)*. Through the NCM module, one can configure various network parameters, including the number of BSs and users, the bandwidth of each spectrum band, the transmission power of the nodes, the simulation time, among others. Experimenters can also specify the pattern following which the blockages are generated and the drone base stations are deployed, e.g., Poisson Point Process and uniform distribution. The NEM module defines the classes for all the network elements, including Network, Ground Base Station, Flying Base Station, Blockage, Links, Interference, among others. These classes have been designed in a hierarchical manner. At the highest level is a general network element class `net_elmt`, which defines the basic network element attributes and operations such as registering an element in the network, specifying the parent and children elements of an element. The DED module provides the discrete network simulation

²It is worth pointing out that another reason for us to conduct simulations based on SimBAG is that we plan to interface SimBAG to software defined experimentation platforms in the future, based on which we can further test FlyTera with testbed experiments.

environment based on the open-source library SimPy [84]. Finally, the CAM module hosts the custom-designed network control algorithms, e.g., the *ESN-Pdt* and *ESN-Opt* algorithms discussed in Section IV. APIs have been provided for all the four modules and the source code of the SimBAG project has been released to the community via GitHub [85].

B. Results and Discussion

A FlyTera network has been designed based on SimBAG with parameters configured through the NCM module. We consider a network area of $200 \times 200 \times 50$ m³. The center frequency is set to 3 GHz, 30 GHz and 300 GHz and the bandwidth is set to 40 MHz, 1 GHz and 10 GHz for the microwave, mmWave and THz bands, respectively. For THz band, the corresponding molecular absorption is set to 0 [67]. The transmission power of BSs is set to 1 W, 250 mW and 20mW for the three bands. The blockages $k \in \mathcal{K}$ are uniformly distributed with minimum and maximum dimensions of 5 m and 20 m, respectively. The network area is discretized into rectangles with length, width and height of 10 m, 10 m and 50 m, respectively. The number of BSs (or users) varies from 1 to 15. The threshold distances d_0^{mc} and d_0^{tz} are set to 100 m and 10 m, respectively. The results are averaged over 50 independent simulations. Next we first evaluate the accuracy and complexity of the echo state learning algorithm, and then analyze the throughput achievable in different spectrum bands.

Accuracy and Complexity. In this experiment we first study the accuracy of the trained ESN module, which has been implemented based on open-source library pyESN [86]. To this end, we consider three mobility models for the UAVs in the training phase: (i) *Random Way Point Model*, (ii) *Horizontal Scanning Model* and (iii) *Diagonal Scanning Model*. For each model, the FBSs move at a speed of 10-15 m/s. The latter two mobility models are illustrated in Fig. 6, and the examples of the resulting trajectories are plotted in Fig. 7, considering two FBSs with the top, middle and bottom subfigures corresponding to the x-, y- and z- coordinates of the FBSs, respectively.

In Figs. 8(a) and (b) we plot the accuracy of *ESN-Pdt* and *ESN-Opt* considering FlyTera with one FBS and 10 users. This experiment is conducted in two steps. (i) *Training Data Generation*: The training data is obtained by moving the FBS

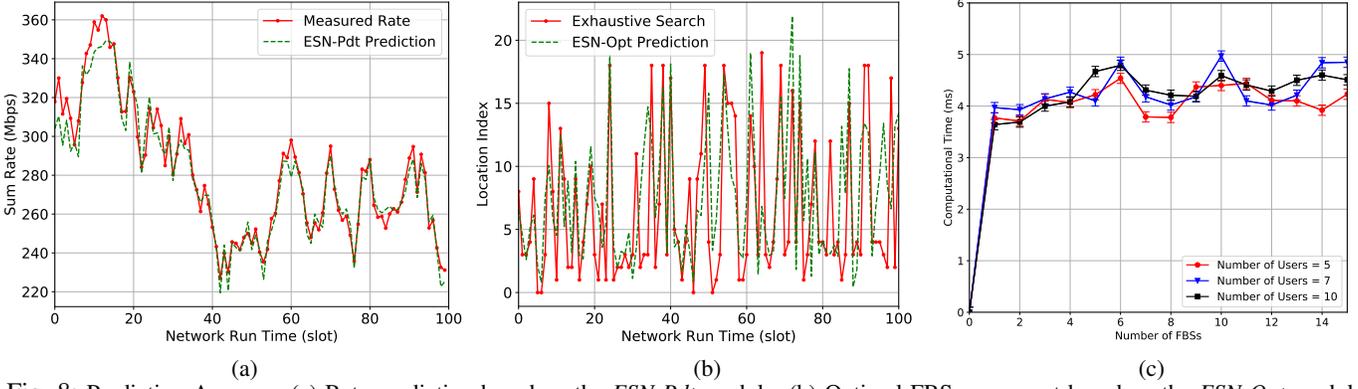


Fig. 8: Prediction Accuracy (a) Rate prediction based on the *ESN-Pdt* module; (b) Optimal FBS movement based on the *ESN-Opt* module; (c) Computational complexity of the *ESN-Opt* module.

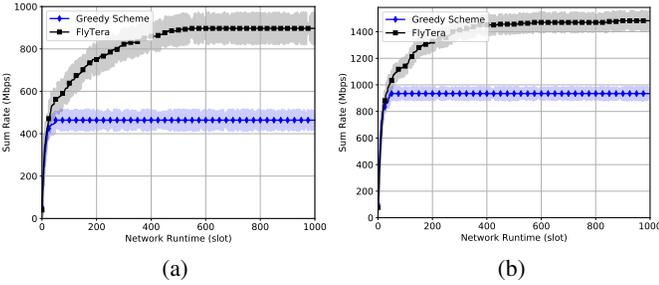


Fig. 9: Performance of FlyTera RL-Ctl Algorithm for (a) Number of FBS = 2; (b) Number of FBS = 5

to a new rectangle in each time slot based on the three mobility models defined earlier for 10000 time slots. (ii) *Training and validation of the model*: The *ESN-Pdt* module is trained with the data from the first 9850 time slots and the remaining data of 150 time slots is used for validation. It can be seen that *ESN-Pdt* is able to effectively predict the rate for the FBS with good accuracy in all the tested time slots. Similar performance can also be achieved by the *ESN-Opt* module as reported in Fig. 8(b).

The testing-phase computational complexity of the ESN learning in terms of time taken for prediction is reported in Fig. 8(c), taking *ESN-Opt* as an example while similar results can be observed for *ESN-Pdt*. The experiments are conducted on a workstation with Intel(R) Core(TM) i7 – 7700K CPU @ 4.20 GHz, memory of 32.0 GB, and 64-bit Windows Operating System. It can be seen that the prediction can be finished in less than 5 ms with different number of users and FBSs. For example, when the number of FBSs varies from 6 to 14, the average computational time is 4.4 ms, 5.0 ms and 4.2 ms for 5, 7, and 10 users, respectively. Therefore, based on echo state learning each FBS is able to predict its optimal movement in each time slot with very low and nearly constant computational complexity in different network settings.

In the training phase, the ESN module solves an RMSE minimization problem (confer (20)) and has a complexity of $\mathcal{O}(N'_{res}{}^2|\mathcal{T}|)$, where $N'_{res} = 1 + N_{inp} + N_{res}$ with N_{inp} being the number of input units, N_{res} the number of reservoir units and $|\mathcal{T}|$ the number of training samples. This may cause high training complexity and low training accuracy in large-scale networks with many UAVs because of the “curse of dimensionality” problem. A possible solution is to divide the

network into a number of subnetworks and train the ESN module for each subnetwork. This will be studied in our future work.

Finally, we study the convergence of the FlyTera Algorithm in Fig. 9. We consider two scenarios with 2 and 5 FBSs and plot the achievable sum rate by averaging over 10 simulations for greedy scheme (i.e., $\epsilon = 0$) and FlyTera scheme (i.e., $\epsilon = 0.1$). The results are reported in Fig. 9 with the confidence interval represented by the shaded region. It can be seen that FlyTera can converge to sum rates significantly higher than that of the greedy scheme. For example, FlyTera achieves sum rate of around 818 Mbps and 1381 Mbps with 2 and 5 FBSs, respectively, which are only 459 Mbps and 818 Mbps for the greedy scheme. This verifies the effectiveness of the FlyTera control scheme.

Performance Analysis. In this experiment we analyze the sum rate performance of the distributed control algorithms proposed in Section IV. We first consider one FBS and the number of users varies from 1 to 15 at step of 2. FlyTera is compared to two benchmark algorithms with fixed BS and randomly moving BS, respectively. For FlyTera, the flight trajectories of the UAVs are controlled by the RL-Ctl module described in Section IV-C. The results are reported in Fig. 10. We found that the network sum rate can be significantly increased by FlyTera, with an average gain of 24% and 40% comparing to random movement and fixed FBS. For example, in the case of 2 users a sum rate of 463 Mbps can be achieved by FlyTera, which is 261 Mbps and 142 Mbps for random movement and fixed FBS, respectively. This experiment further verifies the superiority of the FlyTera control algorithm.

Next, FlyTera is evaluated in three more different scenarios: i) mobile FBS, ii) static FBS with mobile users, and iii) static FBS with static users. Here, for case (ii) we consider that the users move randomly at a speed of 1-3 m/s. In each scenario, FlyTera is compared with six benchmark spectrum access schemes: i) microwave band only, (ii) mmWave band only, iii) THz band only, iv) microwave and mmWave bands, v) microwave and THz bands, and vi) mmWave and THz bands. The results are plotted in Fig. 11. As expected, in all scenarios the sum rate increases with the number of FBSs but at different speeds. For example, in Fig. 11(a) the sum rate increases by 24 Mbps on average by deploying one more FBS in the microwave band, which are 56 Mbps and 52 Mbps for

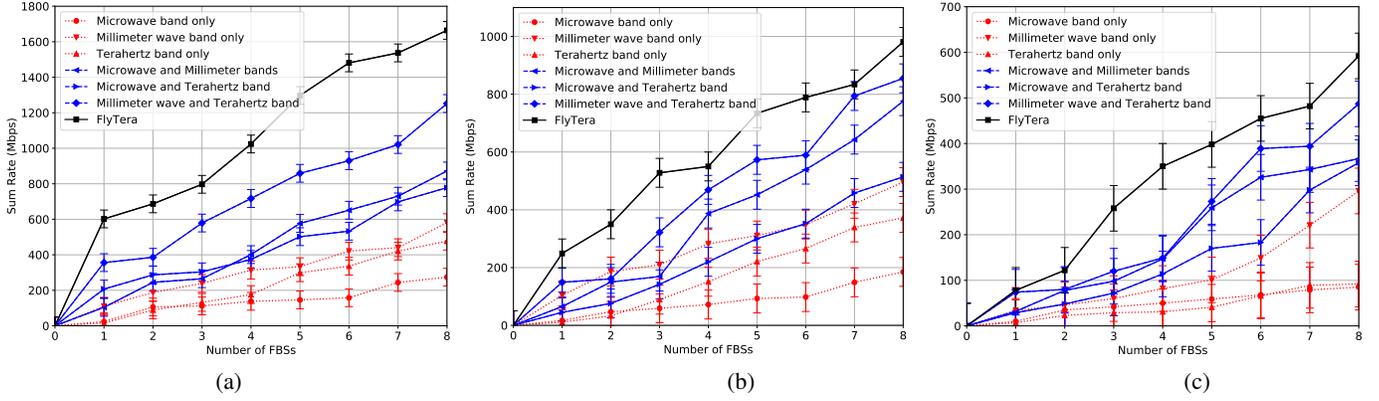


Fig. 11: Sum Rate of Network (a) Mobile FBS, (b) Static FBS with mobile users, and (c) Static FBS with Static users.

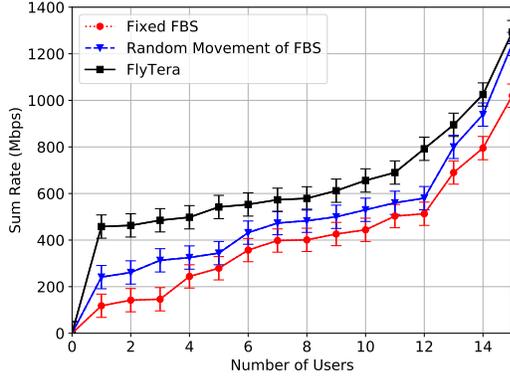


Fig. 10: Sum rate of the network with different number of users.

the mmWave and THz bands, respectively. The corresponding network spectral efficiency gains are 16 bps/Hz/FBS, 2 bps/Hz/FBS and 3×10^{-3} bps/Hz/FBS for the microwave, mmWave and THz bands, respectively. Therefore, although the available bandwidth is much wider, e.g., 10 GHz for THz vs 2 MHz for microwave in this experiment, neither of the mmWave and THz bands can be used alone to achieve orders of magnitude higher network capacity than that of the microwave band, primarily because of the significantly lower spectral efficiency.

From these experiments we also find that, which is a bit surprising, obviously *higher rather than lower* sum rate can be achieved by the THz band in mobile than in static environments. For example, the sum rate is around 500 Mbps and 370 Mbps with 8 mobile FBSs in Figs. 11(a) and 11(b), respectively, while it is only less than 100 Mbps in Fig. 11(c) where both FBSs and users are static. This is without no reasons. While a single THz link can be easily disconnected in blockage-rich environments, the problem can be effectively mitigated in FlyTera by adaptively deploying the flying base stations so that line-of-sight links can be maintained in most time. Therefore, it is important to exploit the mobility gain in future wireless networks in the THz band.

Finally, We find that FlyTera achieves the highest sum rate in all the three tested scenarios, which is 6, 5, and 7 times higher than that using the microwave band only. Particularly, in Fig. 11(a) obviously higher sum rate can be achieved with mobile FBSs than simply adding up the rates of the three single-band cases. For example, in the case of 8

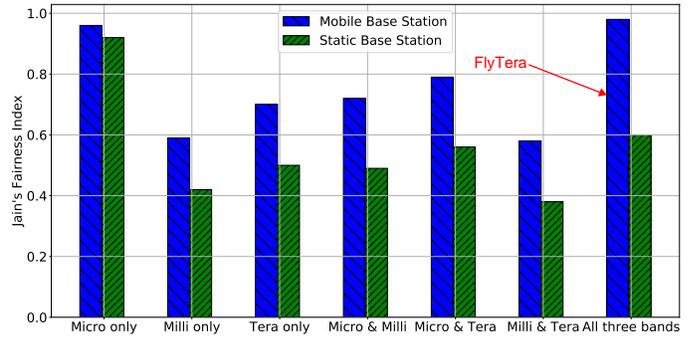


Fig. 12: Rate Allocation Fairness.

FBSs, the sum rate is around 1650 Mbps for FlyTera, while 274 Mbps, 581 Mbps and 477 Mbps for the microwave, mmWave and THz bands with the corresponding sum rate of 1332 Mbps. Similar results can also be observed for the cases using two spectrum bands in Figs. 11(b) and (c). This verifies the effectiveness and importance of joint flight and spectrum access control in FlyTera.

In Fig. 12 we show the fairness results considering mobile and static FBSs and different spectrum access strategies as in Fig. 11. We use Jain's Fairness Index as the measure of the rate allocation fairness. We can see that, in the case of microwave band only, the network achieves the best fairness (0.98), while the data rates of the users are the lowest (Fig. 11). Similarly, when we consider millimeter or terahertz band only, the fairness indices are 0.6 and 0.7 in the case of mobile scenario and only 0.4 and 0.5 in static scenario. The same trend can be observed for the combinations of frequency bands as well. Finally, we can see that with FlyTera the network achieves a fairness index almost the same as that of the microwave band in mobile scenario, while still achieving the highest user rate (Fig. 11). This verifies the capability of FlyTera in achieving a good tradeoff between higher network spectral efficiency and better network coverage.

VI. CONCLUSIONS

In this paper, we have studied the problem of joint flight control and spectrum access in mobile blockage-rich environments in the microwave, mmWave and THz bands. We first provided a mathematical formulation of the FlyTera

control problem, which is shown to be an MINLP problem. Then we designed distributed solution algorithms based on a combination of echo state learning and reinforcement learning. An event-driven simulator called SimBAG has been developed, over which the effectiveness and efficiency of the algorithms are verified and the performance of FlyTera is analyzed through an extensive simulation campaign. It is found that the THz-band wireless networks can significantly benefit from the mobility of FBSs in blockage-rich environments. Future research directions include i) development of link capacity models considering specific transmission strategies and the effects of UAV vibrations; ii) designing multi-timescale ESN-RL algorithms by decomposing the joint spectrum access and flight control problem based on dual method; iii) extending FlyTera by designing more learning algorithms based on, e.g., LSTM, actor-critic methods and deep Q learning; and iv) further verifying the results obtained in this work based on testbed experiments.

REFERENCES

- [1] S. K. Moorthy and Z. Guan, "FlyTera: Echo State Learning for Joint Access and Flight Control in THz-enabled Drone Networks," in *Proc. of IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Como, Italy, June 2020.
- [2] J.-D. Park, S. Kang, and A. M. Niknejad, "A 0.38 THz Fully Integrated Transceiver Utilizing a Quadrature Push-Push Harmonic Circuitry in SiGe BiCMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 10, pp. 2344–2354, Oct. 2012.
- [3] J.-D. Park, S. Kang, S. V. Thyagarajan, E. Alon, and A. M. Niknejad, "A 260 GHz Fully Integrated CMOS Transceiver for Wireless Chip-to-Chip Communication," in *Proc. of Symposium on VLSI Circuits (VLSIC)*, Honolulu, HI, June 2012.
- [4] M. Alibakhshikenari, B. S. Virdee, C. H. See, R. Abd-Alhameed, F. Falcone, and E. Limiti, "A Novel 0.3–0.31 THz GaAs-Based Transceiver with On-Chip Slotted Metamaterial Antenna Based on SIW Technology," in *Proc. of IEEE Asia-Pacific Microwave Conference (APMC)*, Singapore, Dec. 2019.
- [5] C. Wang, B. Lu, C. Lin, Q. Chen, L. Miao, X. Deng, and J. Zhang, "0.34-THz Wireless Link Based on High-Order Modulation for Future Wireless Local Area Network Applications," *IEEE Transactions on Terahertz Science and Technology*, vol. 4, no. 1, pp. 75–85, Jan. 2014.
- [6] B. Blázquez, K. B. Cooper, and N. Llobart, "Time-Delay Multiplexing With Linear Arrays of THz Radar Transceivers," *IEEE Transactions on Terahertz Science and Technology*, vol. 4, no. 2, pp. 232–239, March 2014.
- [7] E. Baştuğ, M. Bennis, M. Médard, and M. Debbah, "Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 110–117, June 2017.
- [8] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Can Terahertz Provide High-Rate Reliable Low Latency Communications for Wireless VR?" *arXiv:2005.00536*, 2020. [Online]. Available: <https://arxiv.org/abs/2005.00536>
- [9] K. Ntontin and C. Verikoukis, "Toward the Performance Enhancement of Microwave Cellular Networks Through THz Links," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 5635–5646, July 2017.
- [10] A. A. Raja, H. Pervaiz, S. A. Hassan, S. Garg, M. S. Hossain, and M. J. Piran, "Coverage Analysis of MmWave and THz-Enabled Aerial and Terrestrial Heterogeneous Networks," *IEEE Transactions on Intelligent Transportation Systems*, June 2021.
- [11] Mumtaz, Shahid et al., "Terahertz Communication for Vehicular Networks," *IEEE Trans. on Vehicular Technology*, vol. 66, no. 7, pp. 5617–5624, July 2017.
- [12] K. M. S. Huq, S. A. Busari, J. Rodriguez, V. Frascolla, W. Bazzi, and D. C. Sicker, "Terahertz-Enabled Wireless System for Beyond-5G Ultra-Fast Networks: A Brief Survey," *IEEE Network*, vol. 33, no. 4, pp. 89–95, July/Aug. 2019.
- [13] H. Elayan, O. Amin, R. M. Shubair, and M. Alouini, "Terahertz Communication: The Opportunities of Wireless Technology Beyond 5G," in *Proc. of International Conference on Advanced Communication Technologies and Networking (CommNet)*, Marrakech, Morocco, April 2018.
- [14] H. Sarrideen, N. Saeed, T. Y. Al-Naffouri, and M. Alouini, "Next Generation Terahertz Communications: A Rendezvous of Sensing, Imaging, and Localization," *IEEE Commun. Magazine*, vol. 58, no. 5, pp. 69–75, 2020.
- [15] G. A. Siles, J. M. Riera, and P. G. del Pino, "Atmospheric Attenuation in Wireless Communication Systems at Millimeter and THz Frequencies [Wireless Corner]," *IEEE Antennas and Propagation Magazine*, vol. 57, no. 1, pp. 48–61, Feb. 2015.
- [16] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-Wideband Waveform Design for Distance-Adaptive Wireless Communications in the Terahertz Band," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 910–922, Feb. 2016.
- [17] C. Han and I. F. Akyildiz, "Distance-Aware Bandwidth-Adaptive Resource Allocation for Wireless Systems in the Terahertz Band," *IEEE Transactions on Terahertz Science and Technology*, vol. 6, no. 4, pp. 541–553, July 2016.
- [18] I. F. Akyildiz and J. M. Jornet, "Realizing Ultra-Massive MIMO (1024x1024) Communication in the (0.06–10) Terahertz Band," *Nano Communications Networks Journal (Elsevier)*, vol. 8, pp. 46–54, June 2016.
- [19] H. Hassanieh, O. Abari, M. Rodriguez, M. Abdelghany, D. Katabi, and P. Indyk, "Fast Millimeter Wave Beam Alignment," in *Proc. of Conference of the ACM Special Interest Group on Data Communication*, Budapest, Hungary, Aug. 2018.
- [20] M. Hashemi, A. Sabharwal, C. E. Koksall, and N. B. Shroff, "Efficient Beam Alignment in Millimeter Wave Systems Using Contextual Bandits," in *Proc. of IEEE ICC*, Honolulu, HI, April 2018.
- [21] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra, "BeamSpy: Enabling Robust 60 GHz Links Under Blockage," in *Proc. of NSDI*, Santa Clara, CA, March 2016.
- [22] A. Alkhateeb, I. Beltagy, and S. Alex, "Machine learning for reliable mmWave systems: Blockage prediction and proactive handoff," in *Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, Nov. 2018.
- [23] I. F. Akyildiz, C. Han, and S. Nie, "Combating the Distance Problem in the Millimeter Wave and Terahertz Frequency Bands," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 102–108, June 2018.
- [24] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter Wave Communication: A Comprehensive Survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 1616–1653, thirdquarter 2018.
- [25] R. Kovalchukov, D. Moltchanov, A. Samuylov, A. Ometov, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Analyzing Effects of Directionality and Random Heights in Drone-Based mmWave Communication," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10064–10069, Oct. 2018.
- [26] W. Xia, V. Semkin, M. Mezzavilla, G. Loianno, and S. Rangan, "Multi-Array Designs for mmWave and Sub-THz Communication to UAVs," in *Proc. of International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Atlanta, GA, May 2020.
- [27] S. Chandrasekharan, A. Al-Hourani, K. Gomez, S. Kandeepan, R. J. Evans, L. Reynaud, and S. Scalise, "Performance Evaluation of LTE and WiFi Technologies in Aerial Networks," in *Proc. of IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, USA, Dec. 2016.
- [28] Y. Shi, R. Enami, J. Wensowitch, and J. Camp, "UABeam: UAV-Based Beamforming System Analysis with In-Field Air-to-Ground Channels," in *Proc. of IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Hong Kong, China, June 2018.
- [29] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, May/June 2020.
- [30] M. Civas, O. Cetinkaya, M. Kuscucu, and O. B. Akanl, "Universal Transceivers: Opportunities and Future Directions for the Internet of Everything (IoE)," 2021. [Online]. Available: <https://arxiv.org/abs/2107.01028>
- [31] M. A. Khan, A. Khan, M. Ahmad, S. Saleem, M. S. Aziz, S. Hussain, and F. M. Khan, "A Study on Flight Time Enhancement of Unmanned Aerial Vehicles (UAVs) Using Supercapacitor-Based Hybrid Electric Propulsion System (HEPS)," *Arabian Journal for Science and Engineering*, vol. 46, no. 2, p. 1179–1198, Sept. 2020.
- [32] J. A. L. Calvo, G. Alirezai, and R. Mathar, "Wireless Powering of Drone-Based MANETs for Disaster Zones," in *Proc. of IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Montreal, QC, Canada, Oct. 2017.

- [33] K. Fujii, K. Higuchi, and J. Rekimoto, "Endless Flyer: A Continuous Flying Drone with Automatic Battery Replacement," in *Proc. of IEEE International Conference on Ubiquitous Intelligence and Computing and IEEE International Conference on Autonomic and Trusted Computing*, Vietri sul Mare, Italy, Dec. 2013.
- [34] J. N. Fadila and P. M. N. S. A. Basid, "Solar Cell-Powered UAVs for Marathon Flights as a Geographic Data Retrieval Tool," in *Proc. of International Conference on Electrical Engineering and Computer Science (ICECOS)*, Batam, Indonesia, Oct. 2019.
- [35] H. Jiang, Y. Niu, B. Ai, Z. Zhong, and S. Mao, "QoS-Aware Bandwidth Allocation and Concurrent Scheduling for Terahertz Wireless Backhaul Networks," *IEEE Access*, vol. 8, pp. 125 814–125 825, July 2020.
- [36] A. Saeed, O. Gurbuz, and M. A. Akkas, "Terahertz Communications at Various Atmospheric Altitudes," *Physical Communication (Elsevier)*, vol. 41, Aug. 2020.
- [37] C. N. Barati, S. Dutta, S. Rangan, and A. Sabharwal, "Energy and Latency of Beamforming Architectures for Initial Access in mmWave Wireless Networks," *Journal of the Indian Institute of Science*, vol. 100, no. 2, p. 281–302, May 2020.
- [38] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Risk-Based Optimization of Virtual Reality over Terahertz Reconfigurable Intelligent Surfaces," in *Proc. of IEEE International Conference on Commun. (ICC)*, Dublin, Ireland, June 2020.
- [39] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C. I., and A. Ghosh, "Millimeter Wave Communications for Future Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, Sept. 2017.
- [40] I. F. Akyildiz, J. M. Jornet, and C. Han, "Terahertz Band: Next Frontier for Wireless Communications," *Physical Communication (Elsevier)*, vol. 12, pp. 16–32, Sept. 2014.
- [41] Z. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Kuang, W. Chen, L. Li, and S. Li, "A Survey on Terahertz Communications," *China Communications*, vol. 16, no. 2, pp. 1–35, Feb. 2019.
- [42] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 836–869, Secondquarter 2018.
- [43] J. Qiu, D. Grace, G. Ding, M. D. Zakaria, and Q. Wu, "Air-Ground Heterogeneous Networks for 5G and Beyond via Integrating High and Low Altitude Platforms," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 140–148, Dec. 2019.
- [44] J. Qiu, D. Grace, G. Ding, J. Yao, and Q. Wu, "Blockchain-Based Secure Spectrum Trading for Unmanned-Aerial-Vehicle-Assisted Cellular Networks: An Operator's Perspective," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 451–466, Jan. 2020.
- [45] W. Yi, Y. Liu, M. ElKashlan, and A. Nallanathan, "Modeling and Coverage Analysis of Downlink UAV Networks with MmWave Communications," in *Proc. of IEEE International Conference on Communications Workshops (ICC Workshops)*, Shanghai, China, May 2019.
- [46] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X. Xia, "3-D Beamforming for Flexible Coverage in Millimeter-Wave UAV Communications," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 837–840, June 2019.
- [47] M. Gapeyenko, V. Petrov, D. Moltchanov, S. Andreev, N. Himayat, and Y. Koucheryavy, "Flexible and Reliable UAV-Assisted Backhaul Operation in 5G mmWave Cellular Networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 11, pp. 2486–2496, Nov. 2018.
- [48] Y. Ke, H. Gao, W. Xu, L. Li, L. Guo, and Z. Feng, "Position Prediction Based Fast Beam Tracking Scheme for Multi-User UAV-mmWave Communications," in *Proc. of IEEE ICC*, Shanghai, China, May 2019.
- [49] Z. Xiao, P. Xia, and X. Xia, "Enabling UAV cellular with millimeter-wave communication: potentials and approaches," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 66–73, May 2016.
- [50] Z. Khosravi, M. Gerasimenko, S. Andreev, and Y. Koucheryavy, "Performance Evaluation of UAV-Assisted mmWave Operation in Mobility-Enabled Urban Deployments," in *Proc. of Int'l Conference on Telecommunications and Signal Processing (TSP)*, Athens, Greece, July 2018.
- [51] Z. Feng, L. Ji, Q. Zhang, and W. Li, "Spectrum Management for MmWave Enabled UAV Swarm Networks: Challenges and Opportunities," *IEEE Communications Magazine*, vol. 57, no. 1, pp. 146–153, Jan. 2019.
- [52] Y. Pan, K. Wang, C. Pan, H. Zhu, and J. Wang, "UAV-Assisted and Intelligent Reflecting Surfaces-Supported Terahertz Communications," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1256–1260, June 2021.
- [53] L. Xu, M. Chen, M. Chen, Z. Yang, C. Chaccour, W. Saad, and C. S. Hong, "Joint Location, Bandwidth and Power Optimization for THz-Enabled UAV Communications," *IEEE Communications Letters*, vol. 25, no. 6, pp. 1984–1988, June 2021.
- [54] S. K. Moorthy and Z. Guan, "Beam Learning in MmWave/THz-Band Drone Networks Under In-Flight Mobility Uncertainties," *IEEE Transactions on Mobile Computing*, accepted for publication, Oct. 2020.
- [55] —, "LeTera: Stochastic Beam Control Through ESN Learning in Terahertz-Band Wireless UAV Networks," in *Proc. of IEEE INFOCOM Workshop on Wireless Communications and Networking in Extreme Environments (WCNEE)*, Toronto, Canada, July 2020.
- [56] L. Zhang, H. Zhao, S. Hou, Z. Zhao, H. Xu, X. Wu, Q. Wu, and R. Zhang, "A Survey on 5G Millimeter Wave Communications for UAV-Assisted Wireless Networks," *IEEE Access*, vol. 7, pp. 117 460–117 504, July 2019.
- [57] J. Hu, H. Zhang, L. Song, Z. Han, and H. V. Poor, "Reinforcement Learning for a Cellular Internet of UAVs: Protocol Design, Trajectory Control, and Resource Management," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 116–123, Feb. 2020.
- [58] S. Yin, S. Zhao, Y. Zhao, and F. R. Yu, "Intelligent Trajectory Design in UAV-Aided Communications With Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8227–8231, Aug. 2019.
- [59] Q. Wang, W. Zhang, Y. Liu, and Y. Liu, "Multi-UAV Dynamic Wireless Networking With Deep Reinforcement Learning," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2243–2246, Dec. 2019.
- [60] X. Zhou, Y. Lin, Y. Tu, S. Mao, and Z. Dou, "Dynamic Channel Allocation for Multi-UAVs: A Deep Reinforcement Learning Approach," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019.
- [61] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 2059–2070, Sept. 2018.
- [62] X. Liu, Y. Liu, and Y. Chen, "Reinforcement Learning in Multiple-UAV Networks: Deployment and Movement Design," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8036–8049, Aug. 2019.
- [63] N. Tafintsev, D. Moltchanov, M. Simsek, S.-P. Yeh, S. Andreev, Y. Koucheryavy, and M. Valkama, "Reinforcement Learning for Improved UAV-based Integrated Access and Backhaul Operation," in *Proc. of IEEE International Conference on Communications Workshops (ICC Workshops)*, Dublin, Ireland, June 2020.
- [64] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y. Liang, and D. I. Kim, "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3133–3174, Fourthquarter 2019.
- [65] P. S. Bithas, E. T. Michailidis, N. Nomikos, D. Vouyioukas, and A. G. Kanatas, "A Survey on Machine-Learning Techniques for UAV-Based Communications," *MDPI Journal of Sensors*, vol. 19, no. 23, pp. 1–39, Nov. 2019.
- [66] J. Jagannath, A. Jagannath, S. Furman, and T. Gwin, "Deep Learning and Reinforcement Learning for Autonomous Unmanned Aerial Systems: Roadmap for Theory to Deployment," *arXiv.org*, 2020. [Online]. Available: <https://arxiv.org/abs/2009.03349>
- [67] J. M. Jornet and I. F. Akyildiz, "Channel Modeling and Capacity Analysis for Electromagnetic Wireless Nanonetworks in the Terahertz Band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, Oct. 2011.
- [68] T. Bai, R. Vaze, and R. W. Heath, "Analysis of Blockage Effects on Urban Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 5070–5083, Sept. 2014.
- [69] S. Yi, Y. Pei, and S. Kalyanaraman, "On the Capacity Improvement of Ad Hoc Wireless Networks Using Directional Antennas," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking Computing*, Annapolis, Maryland, June 2003.
- [70] M. Chen, W. Saad, and C. Yin, "Echo State Learning for Wireless Virtual Reality Resource Allocation in UAV-Enabled LTE-U Networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Kansas City, MO, May 2018.
- [71] M. Chen, O. Semiari, W. Saad, X. Liu, and C. Yin, "Federated Echo State Learning for Minimizing Breaks in Presence in Wireless Virtual Reality Networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 177–191, Jan. 2020.
- [72] C. Gallicchio, A. Micheli, and L. Pedrelli, "Comparison Between DeepESNs and Gated RNNs on Multivariate Time-Series Prediction," *arXiv:1812.11527*, 2018. [Online]. Available: <https://arxiv.org/abs/1812.11527>

- [73] D. Jirak, S. Tietz, H. Ali, and S. Wermter, "Echo State Networks and Long Short-Term Memory for Continuous Gesture Recognition: A Comparative Study," *Cognitive Computation*, Oct. 2020.
- [74] J. Del Ser, I. Laña, E. L. Manibardo, I. Oregi, E. Osaba, J. L. Lobo, M. N. Bilbao, and E. I. Vlahogianni, "Deep Echo State Networks for Short-Term Traffic Forecasting: Performance Comparison and Statistical Assessment," in *Proc. of IEEE International Conference on Intelligent Transportation Systems (ITSC)*, Rhodes, Greece, Sept. 2020.
- [75] H. Peng, C. Chen, C. Lai, L. Wang, and Z. Han, "A Predictive On-Demand Placement of UAV Base Stations Using Echo State Network," in *Proc. of IEEE/CIC International Conference on Communications in China (ICCC)*, Changchun, China, Aug. 2019.
- [76] O. A. Adeleke, "Echo-State Networks for Network Traffic Prediction," in *Proc. of IEEE Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, Oct. 2019.
- [77] X. Yang and F. Zhao, "Echo State Network and Echo State Gaussian Process for Non-Line-of-Sight Target Tracking," *IEEE Systems Journal*, pp. 1–8, April 2020.
- [78] M. Chen, W. Saad, C. Yin, and M. Debbah, "Echo State Networks for Proactive Caching in Cloud-Based Radio Access Networks With Mobile Users," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3520–3535, June 2017.
- [79] M. Chen, W. Saad, and C. Yin, "Echo State Networks for Self-Organizing Resource Allocation in LTE-U With Uplink–Downlink Decoupling," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 3–16, Jan. 2017.
- [80] A. Ferdowsi, M. A. Abd-Elmagid, W. Saad, and H. S. Dhillon, "Neural Combinatorial Deep Reinforcement Learning for Age-Optimal Joint Trajectory and Scheduling Design in UAV-Assisted Networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1250–1265, May 2021.
- [81] Z. Chen, Y. Zhong, X. Ge, and Y. Mia, "An Actor-Critic-Based UAV-BSs Deployment Method for Dynamic Environments," in *Proc. of IEEE International Conference on Communications (ICC)*, Dublin, Ireland, June 2020.
- [82] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction." Cambridge, MA, USA: MIT Press, 1998.
- [83] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory (Classics in Applied Mathematics)*. USA: Society for Industrial and Applied Mathematics, 1999.
- [84] "Simpy." [Online]. Available: <https://pypi.org/project/simpy/>
- [85] <https://github.com/ubwingslab/new-spectrum-technology>.
- [86] "pyESN." [Online]. Available: <https://github.com/cknd/pyESN>



Sabarish Krishna Moorthy is a Ph.D. student in the Department of Electrical Engineering at State University of New York at Buffalo (SUNY Buffalo), NY, USA. He is currently working in the UB WINGS lab under the guidance of Professor Zhangyu Guan. He received his B.E. degree in Electronics and Communication Engineering and M.S. in Electrical Engineering from Velammal Engineering College, Chennai, India and SUNY Buffalo, NY, USA in 2016 and 2018, respectively. His current research interests are in new spectrum technologies and network design automation.

work design automation.



Maxwell McManus is a Ph.D. student in the Department of Electrical Engineering at the State University of New York at Buffalo (SUNY Buffalo), NY, USA. He received the B.S. degree in Electrical Engineering at SUNY Buffalo in 2019, and the M.S. degree in Electrical Engineering at SUNY Buffalo in 2021. He is currently working in the UB Wireless Intelligent Networking and Security (WINGS) lab under the guidance of Dr. Zhangyu Guan. His current research interests are in intelligent wireless network security, digital twin, and spectrum sharing technologies.

technologies.



Zhangyu Guan (SM'11) received the Ph.D. degree in Communication and Information Systems from Shandong University in China in 2010. He is currently an Assistant Professor with the Department of Electrical Engineering at the State University of New York at Buffalo, where he directs the Wireless Intelligent Networking and Security (WINGS) Lab, with research interests in network design automation, new spectrum technologies, and wireless network security. He has served as an Area Editor for Elsevier Journal of Computer Networks since July 2019. He has served as TPC Chair for IEEE INFOCOM Workshop on Wireless Communications and Networking in Extreme Environments (WCNEE) 2020, Student Travel Grants Chair for IEEE Sensor, Mesh and Ad Hoc Communications and Networks (SECON) 2019-2020, Information System (EDAS) Chair for IEEE Consumer Communications Networking Conference (CCNC) 2021. He has also served as TPC member for IEEE INFOCOM 2016-2021, IEEE GLOBECOM 2015-2020, IEEE MASS 2017-2019, IEEE IPCCC 2015-2020, among others.

July 2019. He has served as TPC Chair for IEEE INFOCOM Workshop on Wireless Communications and Networking in Extreme Environments (WCNEE) 2020, Student Travel Grants Chair for IEEE Sensor, Mesh and Ad Hoc Communications and Networks (SECON) 2019-2020, Information System (EDAS) Chair for IEEE Consumer Communications Networking Conference (CCNC) 2021. He has also served as TPC member for IEEE INFOCOM 2016-2021, IEEE GLOBECOM 2015-2020, IEEE MASS 2017-2019, IEEE IPCCC 2015-2020, among others.