

# Unmanned Vehicles in 6G Networks: A Unifying Treatment of Problems, Formulations, and Tools

Winston Hurst, Spilios Evmorfos, Athina Petropulu, Yasamin Mostofi

**Abstract**—Unmanned Vehicles (UVs) functioning as autonomous agents are anticipated to play a crucial role in the 6th Generation of wireless networks. Their seamless integration, cost-effectiveness, and the additional controllability through motion planning make them an attractive deployment option for a wide range of applications. However, despite their potential, the convergence of UVs and wireless systems brings forth numerous challenges that require attention from both academia and industry. This paper then aims to offer a comprehensive overview encompassing the transformative possibilities as well as the significant challenges associated with UV-assisted next-generation wireless communications. Considering the diverse landscape of possible application scenarios, problem formulations, and mathematical tools related to UV-assisted wireless systems, the underlying core theme of this paper is the unification of the problem space, providing a structured framework to understand the use cases, problem formulations, and necessary mathematical tools. Overall, the paper sets forth a clear understanding of how unmanned vehicles can be integrated in the 6G ecosystem, paving the way towards harnessing the full potentials at this intersection.

**Index Terms**—Unmanned Vehicles, 6G, wireless communications, autonomy

## I. INTRODUCTION

6G cellular systems are envisioned to make significant progress in the advancement of wireless communication technology. This has caused much excitement and anticipation in the scientific and high-tech communities, as 6G wireless networks promise unparalleled capabilities that surpass those of 5G systems [1]–[3]. In addition to improving reliability, latency, and capacity by orders of magnitude, 6G will introduce new dimensions to system services, including those related to sensing, artificial intelligence (AI), and positioning, all while responding to societal challenges such as climate change and inequity through an emphasis on sustainability and closing the digital divide [4].

There are a number of inter-related challenges to overcome before realizing the full 6G vision. Some of these are posed directly by the stated network capabilities. How can ubiquitous coverage be achieved in areas where establishing network infrastructure is difficult, such as in maritime environments or in rural areas [5]? Others arise as we consider use cases.

Winston Hurst and Yasamin Mostofi are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93117 USA (email: {winstonhurst, ymostofi}@ece.ucsb.edu). Spilios Evmorfos and Athina P. Petropulu are with the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854 USA (email: se386@scarletmail.rutgers.edu; athinap@soe.rutgers.edu).

This work was supported in part by NSF RI award 2008449 and ARO under Grants W911NF2110071 and W911NF2320103

TABLE I  
LIST OF KEY ABBREVIATIONS.

3GPP	Third Generation Partnership Project
5G	Fifth generation
6G	Sixth generation
AV	Autonomous Vehicle
BCD	Block Coordinate Descent
BS	Base Station
DRL	Deep Reinforcement Learning
HAPS	High Altitude Platform Station
ITU	International Telecoms Union
LEO	Low Earth orbit
LoS	Line-of-sight
MIMO	Multiple-input, multiple output
mmWave	millimeter Wave
MIP	Mixed-integer Program
MILP	Mixed-integer Linear Program
MINLP	Mixed-integer Nonlinear Program
NLoS	Non-LoS
NOMA	Non-Orthogonal Multiple Access
OFDMA	Orthogonal Frequency Division Multiple Access
OMA	Orthogonal Multiple Access
PD-NOMA	Power-Domain NOMA
PRB	Physical Resource Block
QoS	Quality of Service
RL	Reinforcement Learning
RSMA	Rate-Splitting Multiple Access
SCA	Successive Convex Approximation
SDMA	Space division Multiple Access
SOCP	Second-order Cone Program
TBS	Terrestrial Base Station
TDMA	Time Division Multiple Access
THz	Terahertz
UAV	Unmanned Aerial Vehicle
UV	Unmanned Vehicle
UV-BS	UV acting as a BS

For instance, how can throughput, reliability, and latency improvement be guaranteed in dense Internet of Everything (IoE) networks [2], or for extended reality applications [6]. Still more challenges are found as we consider issues associated with specific technologies, such as accounting for the acute attenuation and penetration loss of THz communication [7].

Amidst these challenges, one transformative player emerges in the 6G landscape: *unmanned vehicles*. Autonomous, unmanned vehicles are envisioned to have a crucial impact in shaping the future of 6G cellular systems, enabling them

to overcome the limitations of conventional communication paradigms. For instance, they are anticipated to facilitate novel use cases, enhance network efficiency, improve system intelligence and resilience, and further support the high demands of truly immersive applications.

In this paper, we lay out a critical exploration of the dynamic intersection of unmanned vehicles and 6G cellular systems, aiming to provide a comprehensive and methodical survey that encompasses the underlying challenges, the associated problem formulations, and mathematical tools essential for their integration. **At the core of this research endeavor lies a central theme — the unification of the problem space.** When considering the integration of unmanned vehicles with 6G systems, researchers encounter many different scenarios, several different problem formulations, and an array of mathematical tools. This makes it considerably challenging to comprehend the full spectrum of possibilities in this area and harness them effectively. To address this, this paper takes a fresh perspective by offering a unifying approach. More specifically, it provides a structured framework, with an emphasis on unifying the scenarios, underlying problem formulations, and necessary mathematical tools, in order to enable successful navigation of this complex domain. By doing so, we aim to equip researchers, engineers, and stakeholders with a clear understanding of how unmanned vehicles can be integrated in the 6G ecosystem, which can facilitate the development of innovative solutions that effectively utilize the opportunities presented at this intersection.

We next provide a brief summary of the envisioned capabilities as well as supported applications that 6G can provide, followed by an overview of various types of unmanned vehicles relevant to the 6G ecosystem. This then paves the way for our methodical characterization of the roles UVs can play in the 6G landscape. We end the section with a summary of our contributions.

#### A. 6G: Use Cases and Enabling Technologies

Recent years have seen much discussion and speculation about beyond-5G communication networks, their capabilities, enabling technologies, and supported applications. In June 2023, the International Telecoms Union (ITU) published the IMT-2030 Framework [8], which lays out an official road map for 6G standardization and brings a level of concreteness to the discussion. Core use cases, performance capabilities that support these, and technologies that, in turn, enable the cited performance, are briefly introduced below.

1) *Use Cases and Applications:* The ITU envisions six core usage scenarios for 6G networks, three of which are extensions of 5G capabilities. *Immersive communication* expands upon enhanced mobile broadband (eMBB), with the capacity to support AR and VR experiences [6]. *Massive communication* enhances massive machine-type communication (mMTC) by handling very high traffic loads generated by huge numbers of heterogeneous sensors and devices [10], with envisioned application in smart cities [11] and agriculture [12], to name a few. Moving from 5G to 6G will raise Ultra Reliable and Low-Latency Communication (URLLC) to *Hyper Reliable and Low-Latency Communication* (HRLLC), with performance

KPI	5G (IMT-2020)	6G (IMT-2030)
Reliability	99.999%	99.99999%
Latency	1 ms	0.1 ms
Mobility	500 km/h	1,000 km/h
Connection Density	10 <sup>6</sup> devices/km <sup>2</sup>	10 <sup>8</sup> devices/km <sup>2</sup>
Positioning	--	1 cm
Peak Data Rate	20 Gbps	200 Gbps
User Experienced Data Rate	100 Mbps	500 Mbps
Spectrum Efficiency	--	3X more than 5G
Area Traffic Capacity	10 Mbps/m <sup>2</sup>	50 Mbps/m <sup>2</sup>

Fig. 1. Key performance indicators (KPI) for 5G [9] and 6G [4].

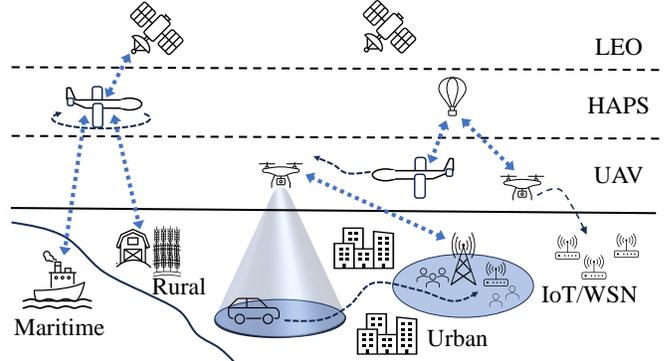


Fig. 2. A high-level view of the various roles UVs play in the 6G ecosystem.

sufficiently robust to support autonomous vehicles in industry 4.0 [3] and remotely controlled medical robotics [13], among other applications.

Three other core use cases are new in 6G. *Integrated sensing and communication* (ISAC) will allow communication signals to also provide information about the physical environment, a feature useful for, e.g., autonomous vehicles navigating through dynamic urban areas [14]. *Ubiquitous connectivity* will close the digital divide, providing cellular and internet access to under-served areas. Finally, the 6G network will empower *integrated artificial intelligence and communication*, facilitating intelligent, data-driven decision-making for many applications, from smart cities [15] to self-driving cars [16].

2) *Enabling the 6G Vision:* Realizing the 6G vision will depend on underlying technical improvements to the network. Figure 1 illustrates several key performance indicators (KPI) and calls out how these improve over current 5G metrics. To make this enhancement possible, a number of important emerging technologies will be employed. We next highlight a few of these.

The *use of high frequency bands* creates greater capacity by permitting larger bandwidths for communication. In 6G, the rollout of mmWave that began in 5G will mature, and sub-THz and THz communication links will come online. However, these spectra bring challenges in a number of dimensions, including hardware design, channel modeling, and rapid signal attenuation [7].

Massive multiple-input, multiple-output (mMIMO) plays a key role in 5G, and multi-antenna techniques will continue to be key enablers in 6G through *extreme MIMO* (eMIMO) [4]. More specifically, higher communication frequencies have smaller wavelengths, and antenna sizes shrink proportionally. This allows very high numbers of them to fit on a single



Fig. 3. 6G: Enabling technologies and enabled applications.

array. Beamsteering can then be used to overcome the fast attenuation of the signal or to provide access to multiple users simultaneously by spatial multiplexing [17].

*Reconfigurable intelligent surfaces (RIS)* or meta-surfaces (MS) [18]–[20] will provide similar beamforming benefits, but in a passive way. These artifacts can be deployed in the environment to dynamically alter the channel quality between nodes by intelligent manipulation of impinging EM fields. By steering signals around large obstacles in the environment, they can ameliorate issues that arise when no line of sight (LoS) path exists between a transmitter and receiver [21].

The high connection density requirements specified in IMT-2030 (see Figure 1) will necessitate *new multi-access schemes*. Emerging candidates, such as Power-Domain Non-Orthogonal Multiple Access (PD-NOMA) [22] and Rate-Splitting Multiple Access (RSMA) [23], are promising methods which can leverage more capacity out of limited communication resources (time, frequency, and space) through intelligent signal processing and interference management.

*Artificial intelligence (AI)* can enhance almost all aspects of the 6G ecosystem, whether for traffic prediction, channel learning, or resource management. These methods are desirable because they can handle the increased complexity of the ecosystem while also allowing for flexible operation to match the dynamic communication environment [24].

Finally, the realization of the 6G vision depends on *diverse network topographies*, as seen in Figure 2. For example, Space-Air-Ground Integrated Networks (SAGIN) and Non-Terrestrial-Networks will be crucial in providing connectivity in scenarios where terrestrial-based cell systems fall short. In this context, *unmanned vehicles (UVs)*, especially UAVs and HAPS, will play a key role in realizing the 6G vision, which is the main motivation for this survey paper. For instance, by leveraging mobility and strategic positioning, they can facilitate dynamic service provisioning, adapting to changing network conditions and traffic patterns, while also improving link quality through direct line-of-sight channels. We next discuss different types of UVs in more detail, before providing a more thorough discussion of their place in 6G and beyond.

## B. Unmanned Vehicles

While the history of unmanned mobile infrastructure in communication systems traces back to the 1940s [25], the utilization of UVs in wireless systems has predominantly been confined to specialized military applications. The advent of the next generation of wireless networking (encompassing Beyond-5G and 6G systems [26]), necessitates adaptable and mobile infrastructure capable of meeting new KPIs. UVs are envisioned to play an integral role in these advanced

wireless systems, unlocking the potential of wireless autonomy by seamlessly integrating components capable of efficient sensing, decision-making, and communication.

Fundamentally, UVs can operate on land, water or in the air. Each medium imposes distinct design considerations. Given their versatility in deployment across various settings, we put a great emphasis on Unmanned Aerial Vehicles (UAVs) in this paper. However, we also address the relevance of Unmanned Ground Vehicles (UGVs) as consumers of 6G services. Notwithstanding their importance, our discussions omit underwater Unmanned Vehicles due to the inherently unique characteristics of communication methods needed in underwater conditions.

6G UAVs, categorized by operational characteristics [27], [28], exhibit altitude-based distinctions. Within this framework, three primary groups emerge: *High-level Aerial Platforms (HAPS)* operate at altitudes ranging from 15km to 25km above ground level, facilitating communication over a radius of 30km and maintaining extended flight duration. *Mid-level Aerial Platforms (MAPS)* navigate altitudes spanning 5km to 15km, covering a communication area with a radius of 5km. These platforms, smaller than HAPS, serve as relays connecting higher-altitude HAPS with lower-altitude ground users or UVs. *Low-level Aerial Platforms (LAPS)* encompass UVs operating at altitudes below 5km above ground level. Generally smaller than counterparts in the other categories, LAPS provide a narrower area of coverage.

UAVs are also categorized in terms of size into distinct subgroups. Micro Aerial Vehicles (MAV) weigh between 50g and 2kg, offering a typical flight range under 10km. Small unmanned Aerial Vehicles (sUAV), within the weight range of 2kg to 50kg, typically feature a 2m wingspan and possess battery capacity supporting a flight range of approximately 15km. Conventional UAVs, spanning weights from 5kg to 150kg, are equipped with battery life sufficient for flight distances ranging from 10km to around 250km. Lastly, Personalized Aerial Vehicles (PAV), designed for large-scale transportation and cargo delivery, exhibit weights from 140kg up to 1500kg and often boast a flight range that exceeds 250km.

Regarding the wing type, UAVs can be equipped with fixed, flapping or rotary wings of varying sizes, depending on the requirements of the flight mission.

The employment of UVs in 6G networks comes with the promise of seamlessly integrating Terrestrial and Non-Terrestrial Networks in the pursuit of connecting both humans and machines [29]. The mobility of UVs enables increased throughput and decreased latency, since the introduced controllable degrees of freedom can be leveraged in order to adjust to fluctuating conditions of network traffic and communication channels. Furthermore, they can assist in achieving ubiquitous and uninterrupted connectivity by dynamically adjusting link quality. Finally, the dynamic motion of UVs promises extension of the network coverage that is being realized by the current immovable infrastructure (static base stations and relay antennas), a direction consistent with the vision of closing the digital divide [30]. In the next section, we provide a framework for understanding the role of UVs in next-generation communication networks.

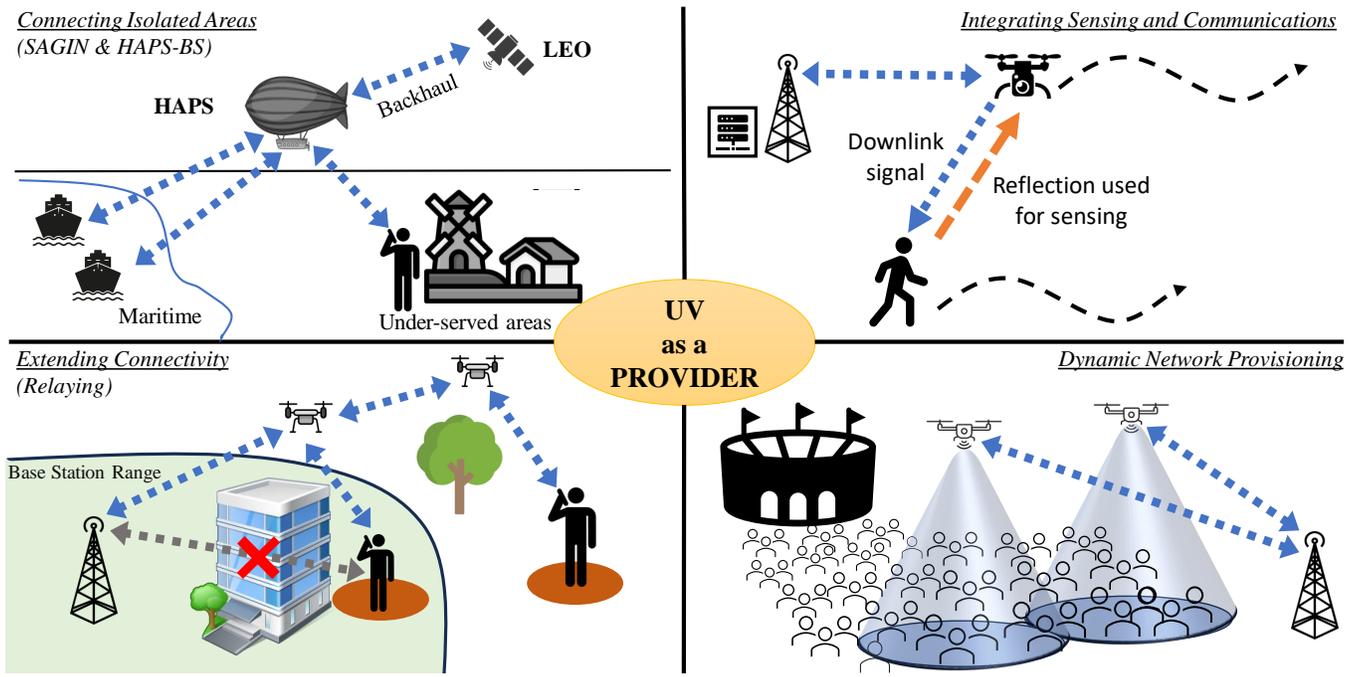


Fig. 4. Sample envisioned applications of UVs as providers in next-generation communication systems. (Top left) As part of a Space-Air-Ground Integrated Network (SAGIN), a HAPS provides connectivity to isolated and under-served areas. (Top right) A UV uses integrated sensing and communication (ISAC) to simultaneously extend connectivity to and further sense end users. (Bottom left) UVs aid a base station by acting as relays. Their elevated position allows them to serve users in dead zones, and by working together, they can extend the area covered by the terrestrial base station (TBS). (Bottom right) A team of UVs provides dynamic network provisioning for a crowd during a large event.

### C. Unmanned Vehicles in the 6G Ecosystem

The use of UVs can both enable and be enabled by 6G networks. **Throughout this paper, we refer to the UV as a provider when its purpose is to aid in the realization of the promises of 6G, and we call it a consumer when it depends on the network’s capabilities to complete a mission.** This duality serves as an organizing principle through the rest of this paper.

1) *UVs as Providers*: Using UVs as providers can address many important challenges. As part of space-air-ground integrated networks (SAIGN), HAPSs can help close the digital divide by obviating expensive build-out of terrestrial infrastructure in rural areas, or provide connectivity to maritime nodes, where infrastructure cannot be built [31]. Furthermore, the elevated nature of aerial UVs makes them well suited to close coverage holes (dead zones), as they frequently experience strong LoS communication channels with ground users. Moreover, their mobility opens the door to real-time network capacity provisioning, either to handle hotspots due to infrequent but acute high-traffic events, to provide temporary solutions when permanent terrestrial infrastructure is offline, or to provision dynamic coverage to a set of mobile users. A sample of these provider roles is shown in Figure 4.

Commercial enterprises have begun making this vision a reality. For example, in 2022, AT&T successfully demonstrated the use of a UAV as a flying 5G base station (BS) [32], while Aalto’s Zephyr HAPS promises direct-to-device 5G connectivity with 5 – 10 ms of latency as soon as 2024 [33]. As recognition of the advantages offered by UVs grows, their use will become more commonplace.

Across the variety of settings in which UVs may be placed as providers, we highlight two general scenarios. First, as UV base stations (UV-BSSs), they must optimize their position and allocation of limited communication resources among multiple users, while mitigating interference with users connecting via other access points. This scenario covers many problems, including the deployment of HAPS and UAVs to provision additional network capacity or take the place of damaged, disabled, or non-existing terrestrial base station (TBS). Second, UVs may be deployed for cooperative communication for relaying using multihop routing or distributed beamforming. This captures problems related to extending the network or covering dead zones.

2) *UVs as Consumers*: On the other hand, UVs as consumers will enjoy robust support in next generation networks, enabling many other exciting applications. 3GPP Release 17, published in 2022, includes specific guidance on incorporating UAVs into the 5G environment [34], with additional studies coming in 2024 as part of Release 18 [35]. Furthermore, the FAA’s Integration Pilot Program, developed to catalyze the incorporation of UAVs into U.S. airspace, published its finding in 2020 [36], and has since been replaced by the BEYOND program, which continues to collect flight data and develop standards, with an eye towards beyond-visual-line-of-sight operation [37].

With standards in development, a number of enterprises have already begun deploying both ground and aerial UVs for a broad range of applications. Well known commercial examples include Alphabet’s Wing [38] and Amazon’s delivery drone program [39], which employ cargo UAVs for last-mile delivery, while Nuro [40] provides a similar ser-

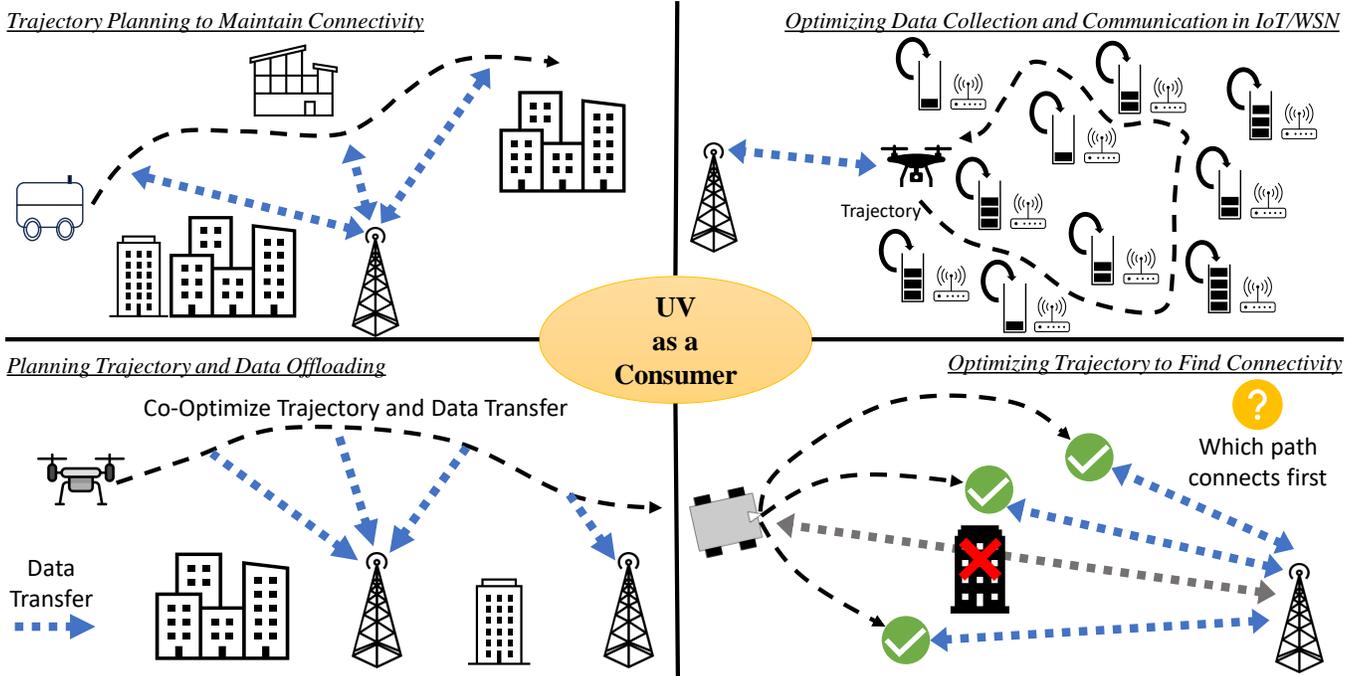


Fig. 5. Sample common tasks for UV operation in 6G networks. (Top left) A UV navigates to a destination while maintaining connectivity. (Top right) A UV collects data from a dynamic WSN. (Bottom left) A UV must jointly plan its trajectory and transmission power/communication rate while offloading data to a remote station. (Bottom right) A UV, initially disconnected from the network, seeks to reconnect as quickly as possible.

vice with autonomous ground vehicles. Autonomous vehicles have also found applications in warehouse management, with Amazon’s Proteus and Sequoia platforms demonstrating the possibilities in this area [41]. Future industrial applications include precision agriculture and facility surveillance [42]. Furthermore, Tesla [43], Waymo [44], and Uber [45] all have self-driving cars in various stages of development, and the vision of fully autonomous automobiles has taken root across the industry. In all these cases, communication can facilitate mission completion by allowing for command and control, information sharing, and access to edge computing services.

For optimal performance, UVs must jointly optimize their trajectories and communication as they complete their missions, with a few key problem scenarios deserving attention. Most fundamentally, UV users must maintain connectivity for command and control, or, if connectivity is lost, they must reconnect as soon as possible. They may additionally need to schedule wireless offloading of data collected from, e.g., onboard cameras, IoT devices, or wireless sensor networks, while accounting for the quality of the underlying communication channel. The data collection itself may involve deciding on the best way to visit a number of sites while taking into account the timeliness of data transfer back to a remote base station. Examples of these problems are illustrated in Figure 5.

#### D. Our contributions

Using the mobility of unmanned vehicles to enable new forms of connectivity is not a new area, and dates back to work as early as 2008. For instance, in [46] the term *communication-aware motion planning* was coined to refer to path planning that considers its impact on communication parameters, while [47] showed how a UV can actively plan

its path to enforce new forms of connectivity. This motivated subsequent work since then that showcased the potentials of unmanned vehicles to enable new forms of connectivity [48]. While these explorations were for a general system, the possibilities and challenges offered by the 6G vision created great new interest in this area since 2019.

Along this line, the past few years have produced a number of valuable surveys and tutorials on UVs in communication systems [49]–[56]. Some of these focus on particular topics, such as machine learning [52], IoT networks [49], or path planning [56], while others provide a high-level perspective of the potentials and challenges in this space. Moreover, most such work are written from a 5G perspective. In this context, this paper takes a unique perspective by focusing on the optimization problems that arise when planning the operation of UVs in 6G communication networks. Its primary contribution lies in the unification of the problem space, providing a structured framework to understand the use cases, underlying problem formulations, and the needed mathematical tools. Overall, the paper aims to furnish the readers with a systematic understanding of the potentials and challenges offered by the integration of UVs into the 6G communication systems.

The rest of the paper is organized as follows. In Section II, we introduce the two key types of optimization variables — those related to communication resource allocation and UV mobility — and discuss key modeling considerations for each in the context of 6G. Section III discusses the aspects of these systems that make optimization mathematically challenging and provides a set of key tools that can be used to navigate this space. With this groundwork laid, Section IV then takes a unifying perspective of the problem formulation space, methodically categorizing the large array of potential scenarios

into a set of fundamental problems, all the while bringing key insights into relevant sub-problems. Finally, Section V gives other important considerations for the optimization of UV operation, which is followed by concluding remarks in Section VI. We note that Sections II and III are written so that readers familiar with their content may choose to move directly to Section IV.

## II. KEY CONCEPTS AND MODELING

The previous section highlights the many roles UVs will play in next-generation communication systems, but the realization of UVs' full potential depends on intelligent, optimized deployment of these assets in the network. In Section IV, we provide a unifying and structured treatment of the corresponding problem space. This section then identifies core components of these systems and presents relevant models needed for tackling the problem space. Throughout, we identify common types of constraints and performance metrics, and we end the section with a high-level problem formulation.

### A. Common Core Components

To identify common system components, consider  $N_b$ , a bound on the number of bits that can successfully be sent between a transmitter and a receiver over a given time period,  $\Delta t$ . From Shannon's channel capacity for additive white Gaussian noise (AWGN) [57], this can be expressed as

$$N_b = \Delta t B \log_2(1 + \text{SINR}), \quad \text{SINR} = P|H|^2/(\sigma^2 + I) \quad (1)$$

where  $B$  is the channel bandwidth, SINR is the signal-to-interference-plus-noise ratio,  $P$  is the transmit power,  $|H|$  is the channel gain,  $\sigma^2$  is the noise power at the receiver, and  $I$  is interference power that arises from other communication links using the same physical resource block (PRB).

This bound indicates the two avenues through which UV operation impacts communication. First,  $N_b$  depends on the allocation of wireless resources: the time spent communicating over the channel ( $\Delta t$ ), the transmit power ( $P$ ), and the bandwidth of the channel ( $B$ ). In general, these resources may need to be allocated among multiple users simultaneously, and the interference,  $I$ , may depend on allocation to other links in the network. Second,  $N_b$  is a function of the communication channel,  $H$ , which, while generally considered exogenous, is controllable to a degree through the mobility of the UV.

Based on the discussion above, we identify three general components of these systems which merit attention. First, the communication channel affects the quality of communication over a link. Second, the allocation of communication resource across over multiple users, together with the channel, determines the communication performance. Third, mobility gives the UV some control over the channel while also allowing for the completion of other mission-specific objectives. We next discuss each of these high-level components, with an emphasis on considerations for 6G and beyond.

### B. Communication Channel Models

The channel plays a central role in the performance of the system, and thus appropriate modeling is very important.

However, channels display complex spatio-temporal variation, making them challenging to model and predict. We next discuss several key modeling concepts and then move to considerations for 6G and beyond.

1) *Statistical Channel Models*: The spatial dynamics of channels are generally modeled as the superposition of three components. *Path loss* describes the large-scale attenuation of the signal power as it propagates through space. *Shadowing* captures the impact of large-scale blockages which create Non-LoS (NLoS) communication links. Finally, *multi-path* or *small-scale fading* models the small-scale variation due to multiple copies of the signal arriving at a receiver via different paths. When the multipath delay spread is relatively small, a general model has the form

$$|H|^2 = H_{\text{PL},0} d^{-\alpha_{\text{PL}}} \alpha_{\text{SH}}^2 \gamma_{\text{MP}} \quad (2)$$

where  $H_{\text{pl},0}$  is the *path loss intercept* (i.e., the path loss at a reference distance of 1 m),  $\alpha_{\text{PL}}$  is the *path loss exponent*,  $d$  is the distance between the transmitter and the receiver,  $\alpha_{\text{SH}}^2$  is the shadowing power, and  $\gamma_{\text{mp}}$  is the multipath effect. While the path loss intercept and exponent are modeled as scalar values, shadowing,  $\alpha_{\text{SH}}^2$ , is generally modeled as a spatially-correlated log-normal random variable, and multipath effects can be modeled with Nakagami, Rayleigh, Rician, or log-normal distributions [58], [59]. The values of the path loss intercept and path loss exponent parameters, as well as the parameters for the shadowing and multipath distributions, depend on the carrier frequency and environmental factors. For the case of wideband channels where delay spread needs to be considered, the same modeling principles can be extended over the delay dimension, resulting in the overall power-delay profile of the channel.

2) *Channels for UAVs and HAPSs*: For UVs deployed at high altitudes, the shadowing and multipath components may be omitted to model scenarios of unimpeded propagation between nodes, both for air-to-air links between UVs and for air-to-ground links when the height of a HAPS or UAV results in a dominant line-of-sight (LoS) link [60].

For air-to-ground channels, shadowing is often accounted for by using different path loss intercepts and exponents depending on whether a dominant LoS path exists between the transmitter and receiver:

$$|H|^2 = \begin{cases} H_{\text{PL}, \text{L},0} \|q_i - q_j\|_2^{-\alpha_{\text{PL}, \text{L}}} \text{ w/ prob. } p_{\text{LoS}} \\ H_{\text{PL}, \text{N},0} \|q_i - q_j\|_2^{-\alpha_{\text{PL}, \text{N}}} \text{ w/ prob. } (1 - p_{\text{LoS}}) \end{cases} \quad (3)$$

where  $H_{\text{PL}, \text{L},0}$ ,  $\alpha_{\text{PL}, \text{L}}$  and  $H_{\text{PL}, \text{N},0}$ ,  $\alpha_{\text{PL}, \text{N}}$  are the path loss intercept and exponent for the LoS and NLoS conditions, respectively, and  $p_{\text{LoS}}$  is the probability of the existence of the LoS path.

Three common models for calculating  $p_{\text{LoS}}$  appear in the literature. A method used by the ITU relies on three parameters to statistically characterize the probability of a LoS channel existing between a UAV and a ground node [61]. The parameters of this model are  $\beta$ , the number of buildings per square kilometer,  $\alpha$ , the fraction of area covered by buildings, and

$\kappa$ , the parameter for a Rayleigh distribution describing the building heights. The formula is given by

$$P_{\text{LoS}} = \prod_{n=0}^{\max(0, o-1)} \left[ 1 - \exp \left( -\frac{\left( h_{\text{gr}} + \frac{\Delta h(n+0.5)}{o} \right)^2}{2\kappa^2} \right) \right],$$

where  $h_{\text{gr}}$  is the height of the ground node,  $o = \lfloor \sqrt{\alpha\beta} \rfloor$  with  $\lfloor \cdot \rfloor$  the floor operator, and  $\Delta h$  is the difference in height between the UAV and the ground node.

Alternatively, 3GPP has produced recommendations TR36.777 [62] and TR36.873 [63], which calculate  $P_{\text{LoS}}$  as:

$$P_{\text{LoS}} = \begin{cases} 1, & \text{if } d_{\text{horiz}} \leq d_1 \text{ or } h_u \geq 100 \\ \frac{d_1}{d_{\text{horiz}}} + \left( 1 - \frac{d_1}{d_{\text{horiz}}} \right) \exp \left( \frac{-d_{\text{horiz}}}{d_2} \right), & \text{if } d_{\text{horiz}} > d_1 \end{cases},$$

where  $h_u$  is the height of the UAV,  $d_{\text{horiz}}$  is the horizontal distance between the UAV and the ground node,  $d_1 = \max\{460\log_{10}h_u - 700, 18\}$ , and  $d_2 = 4300\log_{10}(h_u) - 3800$ . Importantly, this model holds for  $22.5 \text{ m} < h_u \leq 300 \text{ m}$ .

Finally, work in [64], [65], based on the ITU model, finds that in many scenarios, the probability of a LoS channel is modeled well as

$$P_{\text{LoS}} = \frac{1}{1 + b_1 \exp(-b_2(\theta - a))}, \quad (4)$$

where  $b_1$  and  $b_2$  are environment parameters, and  $\theta$  is the UAV's angle of elevation with respect to the ground node.

Additional details on air-to-ground communication channels may be found in [66], [67].

3) *Communication Channels at mmWave and Beyond:* While the modeling discussed thus far suffices for a wide range of frequencies, modeling 6G frequencies (mmWave and THz) requires additional considerations. Modeling these channels continues to be an area of active research, with details found in [68], [69].

*Acute Signal Attenuation:* mmWave and THz communication signals attenuate much more quickly as they travel through space, due to the physics of electromagnetic propagation as well as increased impacts of atmospheric absorption, necessitating the use of beamforming for energy-efficient communication. Furthermore, signals at these frequencies do not propagate well through most solid material due to acute penetration loss. Consequently, the existence of a LoS path between the transmitter and receiver becomes more important, though technologies such as RIS may mitigate this challenge by steering around the blockage.

*Sparse Multipath* In part due to acute signal attenuation, 6G frequencies tend to be multipath sparse, so the channels tends to be dominated by path loss and shadowing.

4) *Temporal Channel Variation:* Importantly, communication channels may vary across times as well as space, due to changes in the environment. To account for this, a block-fading model assumes the channel is temporally stationary over a coherence time of length  $\Delta_C$  (see, e.g., [70]). Alternatively, the shadowing power over both space and time may be modeled with a Gaussian process with an autocorrelation function

$$R(\alpha_{\text{SH}}(q, t), \alpha_{\text{SH}}(q', t')) = \alpha_{\text{SH}}^2 e^{-\frac{\|q-q'\|_2}{d_c} - \frac{|t-t'|}{t_c}}, \quad (5)$$

where  $d_c$  and  $t_c$  give the rate of decorrelation in space and time, respectively [71].

5) *Channel Prediction:* For fixed transmitters and receivers, there are many established techniques to estimate the channel gain after receiving a packet. However, for trajectory planning purposes, it is often important to know the channel gain at points not yet visited by the UV. The UV may employ a statistical modeling framework for channel prediction across the workspace. For example, given a few prior channel measurements, [59] proposes a Gaussian-process based predictor which gives the probability distribution over the channel gain values at any unvisited location in the workspace. More recently, [72] uses a convolutional deep neural network to predict channel values across the workspace with only a few prior measurements and without knowledge of the transmitter location. In [73], the authors demonstrate how the mobility of a UV can be used to facilitate full ray makeup channel prediction, enabling the prediction of the magnitude, delay and angle of arrival of all the incoming paths at unvisited locations.

[74] extends the Gaussian process model across space and time for predicting spatio-temporal variations. The work of [75] proposes a distributed stochastic optimization method in order to predict the SNR at future time steps for mobile relay motion control, assuming that the log-magnitude of the channel evolves as a Gaussian process that is correlated across time and space along the lines of (5). The work by [76] proposes a correlated bandits approach in order to predict the spatio-temporally correlated channels.

6) *Ray Tracing:* Ray tracing models communication channels by tracing the path of electromagnetic radiation from the transmitter to the receiver and explicitly accounting for phenomena such as reflection, diffraction, refraction, and scattering by splitting, attenuating, and redirecting the rays. This approach requires detailed information about the environment, and for all but the simplest cases needs computationally-intensive simulation. As a result, these methods have generally been used in offline settings, with desktop applications such as Remcom's Wireless InSite [77] capable of producing high-fidelity simulated channels. However, recent developments in mobile hardware, such as Qualcomm's Snapdragon 8 Gen 2 [78], have raised the prospect of real-time ray tracing on small, energy-limited devices. This, in conjunction with real-time data from UV-mounted vision systems, may eventually permit this approach for channel modeling on the fly.

7) *Realistic Channel Environments:* Regardless of which channel model is used, good modeling and validation require grounding in real-world measurement campaigns. Standards and recommendations written by bodies such as the ITU and 3GPP are good resources for finding realistic parameters for statistical channel modeling [61]–[63], as are academic works which take on the challenge of measurement campaigns, several of which are summarized in [67].

Alternatively, realistic channel data may be synthesized. For example, [79] generates statistically realistic channel spatial variations using a Gaussian process model for shadowing (in dB domain), with an exponential correlation function, as discussed earlier, while also accommodating multipath effects if needed, based on the work in [80], and [81] provides

sophisticated channel simulation for 5G and 6G mmWave and sub-THz channels in environments studied by 3GPP. Similarly, the work of [82] (as outlined in Theorem 2) shows how to synthesize channels that have the spatio-temporal characteristics of (5) based on a stable order-1 vector autoregression, with a corresponding software package found in [83]. As noted earlier, when sufficient environmental information is available, ray tracing software may also be used.

8) *Model-Free Channels*: While we have focused on analytical or simulated models, data-driven, model-free approaches offer an interesting avenue to capture channel characteristics, particularly in the context of Reinforcement Learning [84]. A key challenge here lies in undertaking measurement campaigns large enough to capture channel variations across both space and time.

### C. Communication: Interference Management, and Metrics

For a single communication link, the allocation of limited power and bandwidth over a finite time, together with the channel, determine the performance of the communication system, as indicated by the bound in (1). When multiple links are considered, the picture becomes more complex due to the possibility of interference and the need to aggregate the performance over each link. We next discuss interference management via multiple access schemes before reviewing important metrics of the communication system's performance.

1) *Multiple Access Schemes*: Multi-access schemes give structure to the allocation process by defining rules which help mitigate interference between users connecting to the same BS. Traditionally, orthogonal multiple access (OMA) schemes have been used to ensure that no two users send or receive data on the same PRB by splitting either time (time-division multiplexing (TDM)), or bandwidth (orthogonal frequency division multiplexing (OFDM)) into equal, non-overlapping partitions of a predetermined size and assigning a single user to each.

To achieve the objectives of 6G, physical resources will need to be used more efficiently, leading to increased interest in non-orthogonal multiple access schemes, which allow multiple channels to occupy the same PRB. For example, massive MIMO, a key enabler of 5G, allows for space-division multiple access (SDMA) by creating strong antenna gains in multiple directions simultaneously [17]. In 6G, *ultra-massive* MIMO [6] or *extreme* MIMO [4] will continue to develop this approach. Alternatively, PRBs can be reused by employing intelligent signal processing and interference management with techniques such as power-domain NOMA (PD-NOMA) [22] and rate-splitting multiple access (RSMA) [23]. Table II provides an overview of these schemes with associated resource constraints and inter-cell interference expressions.

2) *Communication Metrics*: In IMT-2030, the ITU has set out a variety of metrics, both new and long-standing, to measure the performance of 6G systems, as we review next.

Fundamentally, the SINR of individual communication links determines the performance of the network. The SINR in turn determines the maximum achievable spectral efficiency,  $r^{\max} = \log_2(1 + \text{SINR})$ , and by extension the maximum

achievable communication rate. The performance may also be measured in terms of energy efficiency, *i.e.*, bps/W.

On the other hand, by modeling data generation at an end user as an  $M/G/1$  queue, the latency across a single link is

$$l = \frac{\lambda}{2(Br^{\max}\nu)((Br^{\max}/\nu) - \lambda)} + \frac{\nu}{Br^{\max}},$$

where  $\nu$  is the packet size [85] and  $\lambda$  is the data generation rate for user  $u$  in bps.

Another key communication metric is the bit-error rate (BER), which is the ratio of the number of bit errors to the total number of transmitted bits. The BER may be calculated in various ways depending on factors such as underlying modulation, coding scheme, and channel characteristics [58]. Assuming an M-QAM communication constellation with spectral efficiency  $r$ , the reliability in terms of the bit error rate is well approximated by

$$p_b \approx 0.2 \exp(-1.5 \text{SNR} / (2^r - 1)).$$

for certain SNR regimes [58].

Other important metrics include coverage, outage probability, and the number of handoffs experienced by a user, *i.e.*, the number of times the user's BS association changes, while secrecy rate provides a way of evaluating the security and privacy of the system [86].

### D. UV Mobility: Kinematics and Energy Models

So far, we have discussed key communication concepts and models. We now address another important facet, the positioning and movement of UVs. Mobility supports UVs both as providers and consumers in the network by giving some degree of control over the communication channel, while also allowing for the completion of mission-specific tasks such as data harvesting or automated delivery. We next briefly present key considerations for modeling the mobility of UVs including kinematic and energy consumption models.

1) *Dynamics and Kinematic Constraints*: A basic model treats the UV as a single point,  $q$ , constrained to lie in some region  $\mathcal{Q} \in \mathbb{R}^3$ , which may account for obstacles or no-fly zones. In the multi-UV case, a minimum distance constraint between UVs may enforce collision avoidance.

A basic second-order dynamic system provides a basic model of mobility, expressed as

$$\dot{q} = v, \quad \dot{v} = u, \quad (6)$$

where the dimensions of velocity,  $v$ , and control input (acceleration),  $u$ , match the dimension of  $q$ . When planning a 2D trajectory, either as a ground vehicle or when flying at a fixed altitude, the UV's velocity,  $v$ , is also constrained to be in some range, *i.e.*,  $v_{\min} \leq \|v\| \leq v_{\max}$ , with the minimum and maximum values dependent on the specifics of the UV. For ground vehicles and rotary-wing UAVs,  $v_{\min} = 0$ , while fixed-wing UAVs will have some non-zero minimum flight speed. For 3D trajectories, the velocity constraint may be further decomposed into two separate vertical and horizontal maximum velocities, as for many UAVs, the mechanics of moving vertically differ significantly from those of moving horizontally [87], [88].

TABLE II

IMPORTANT MULTI-ACCESS COMMUNICATION SCHEMES WITH RELATED CONSTRAINTS AND INTERFERENCE VALUES. HERE,  $P_u$  IS THE POWER ALLOCATED TO USER  $u$ , AND  $P_{\text{MAX}}$  IS THE MAXIMUM AVAILABLE TRANSMIT POWER.  $s_{uc}$  AND  $f_{uc}$  ARE BINARY VARIABLES INDICATING WHETHER USER  $u$  IS USING TIME SLOT (TDM) OR SUB-CHANNEL (OFDM)  $c$ , RESPECTIVELY.  $\vec{w}_u$  IS THE BEAMFORMING VECTOR FOR USER  $u$ , AND  $\vec{h}_u^H$  IS THE HERMITIAN TRANSPOSE OF THE CHANNEL VECTOR  $\vec{h}_u = [H_{u,1}, \dots, H_{u,j}, \dots, H_{u,N_t}]$  WHERE  $H_{u,j}$  IS THE CHANNEL BETWEEN THE  $j^{\text{TH}}$  ANTENNA AND THE USER, AND  $N_t$  IS THE NUMBER OF TRANSMIT ANTENNAS. FOR PD-NOMA, THE USERS ARE SORTED BY THE AMOUNT OF POWER ALLOCATED TO EACH AND INDEXED IN ASCENDING ORDER. FOR RSMA,  $P_{\text{CMN}}$  IS THE POWER ALLOCATED TO THE COMMON COMPONENT AND  $P_u$  REFERS TO THE PRIVATE COMPONENT FOR EACH USER.

Scheme	Channels Defined by	BS DL Power Constraint	Channel Constraints	DL Intra-cell Interference
TDM	Orthogonal time allocation	$P_u \leq P_{\text{MAX}}$	$\sum_{u \in \mathcal{U}} s_{u,c} \leq 1$	None
OFDM	Orthogonal frequency blocks	$\sum_{u \in \mathcal{U}} P_u \leq P_{\text{MAX}}$	$\sum_{u \in \mathcal{U}} f_{u,c} \leq 1$	None
SDMA	Quasi-orthogonal spatial beams	$\sum_{u \in \mathcal{U}} \ \vec{w}_u\ _2^2 \leq P_{\text{MAX}}$	Limited by number of antennas	$I_u = \sum_{u' \in \mathcal{U}, u' \neq u}  \vec{h}_u^H \vec{w}_{u'} ^2$
PD-NOMA	Channel reuse	$\sum_{u \in \mathcal{U}} P_u \leq P_{\text{MAX}}$	-	$I_{u_i} = \sum_{j=i+1}^n P_{u_j}  H_{u_i} ^2$
RSMA	Partial channel reuse	$P_{\text{cmn}} + \sum_{u \in \mathcal{U}} P_u \leq P_{\text{MAX}}$	-	See [23]

Acceleration should be similarly bounded to capture realistic kinematic constraints.

This basic model has the advantage of simplicity and serves when planning for rotary-wing UAVs or, more generally, when the UV has a small turn radius compared to the scale on which the planning occurs. At other times, however, more detailed kinematics are important. For many ground vehicles, a basic car model offers a simple yet descriptive representation [89]. This differential model includes not only the position of the vehicle,  $q$ , but the orientation  $\theta_o$  as well. The system takes as input a velocity,  $v$  and steering angle,  $\phi$ , and the dynamics are

$$\dot{q} = v[\cos\theta_o, \sin\theta_o], \dot{\theta}_o = \phi_o, \quad (7)$$

with  $\dot{q} = [\dot{q}_x, \dot{q}_y]$  denoting the velocity in the x- and y-coordinates. Generally, the steering angle is constrained, i.e.,  $|\phi_o| \leq \phi_{o,\text{MAX}}$ , and the minimum turning radius is given by  $L/\tan\phi_{o,\text{MAX}}$ , with  $L$  as the distance between the two axles. If  $v$  is allowed to be negative, the vehicle may go in reverse, but when the velocity is constrained so that  $v \in \{0, 1\}$ , we have the well-studied Dubins model [90].

The kinematics of fixed-wing aircraft are quite complex, but a simple extension of the model above gives an elementary yet useful set of dynamics. In particular, the dynamics at a fixed flying altitude and constant velocity can be modeled directly as a Dubins model, while any changes in altitude can be captured by the independent linear system  $\dot{q}_z = v_z$ , for a vertical velocity  $v_z$  [89].

The above models give a more accurate sense of a UVs movement capabilities than the simple point model first presented. However, some scenarios, such as when studying air-frame occlusion or beam-misalignment, require rigorous understanding of UV kinematics [67], [91]. In these cases, appropriate models must be found using established techniques from robotics.

2) *UV Energy Models*: UVs generally have access to limited energy, so understanding the energy requirements for a proposed trajectory is critical to successful operation. As with kinematics, exact expressions for energy consumption are tied to the specifics of the UV under consideration, but some general models capture important features for different types of UVs. For example, [92] develops an empirical model for a pioneer ground robot given by  $P(v) = \kappa_1 v + \kappa_2$ , where  $\kappa_1$  and  $\kappa_2$  are determined by the load and mechanics of the robot. Alternatively, [93] gives the instantaneous power consumption

for a rotary-wing UAV as a function of velocity with the following formula:

$$P(v) = \underbrace{P_0 \left( 1 + \frac{3v^2}{U_{\text{tip}}^2} \right)}_{\text{blade profile}} + \underbrace{P_i \left( \sqrt{1 + \frac{v^4}{4v_0^4}} - \frac{v^2}{2v_0^2} \right)}_{\text{induced}} + \underbrace{\frac{1}{2} d_0 \rho s_r A v^3}_{\text{parasite}} \quad (8)$$

where,  $v$  is the velocity,  $P_0$  and  $P_i$  are constants,  $U_{\text{tip}}$  is the speed of the tip of the rotor blade,  $v_0$  is the mean rotor induced velocity in hover,  $d_0$  is the fuselage drag ratio,  $s_r$  is the rotor solidity,  $\rho$  is the air density, and  $A$  is the rotor disk area. Importantly, this model is not convex in  $v$ , but a convex approximation fits well outside the low-velocity regime.

Finally, for a fixed-wing UV flying at a fixed altitude, [94] derives an energy model which accounts for drag, lift, and thrust. For the velocity vector  $\mathbf{v} \in \mathbb{R}^2$  and its time derivative,  $\dot{\mathbf{v}}$ , the expression is given as:

$$P(\mathbf{v}, \dot{\mathbf{v}}) = c_1 \|\mathbf{v}\|^3 + \frac{c_2}{\|\mathbf{v}\|} \left( 1 + \frac{\|\dot{\mathbf{v}}\|^2 - \frac{\dot{\mathbf{v}}^T \mathbf{v}}{\|\mathbf{v}\|^2}}{g^2} \right) + \frac{m \|\dot{\mathbf{v}}\|}{4}, \quad (9)$$

where  $m$  is the mass of the UV,  $g$  is acceleration due to gravity, and  $c_1$  and  $c_2$  are constants determined by other parameters of the UV air frame. The  $1/\|\mathbf{v}\|$  and  $\|\mathbf{v}\|^3$  terms capture high power consumption at both very low and very high velocities.

### E. A High Level Problem Formulation

To conclude, we present a high-level problem formulation consisting of an objective — which may account for the communication performance/cost, motion costs, and other mission objectives — and constraints which can be motivated by communication, motion, or mission-specific considerations. The decision variables can be broadly categorized as either communication- or trajectory-related, with the possible addition of mission-related variables representing concepts such as sensing. Let  $\mathcal{R}$ ,  $\mathcal{T}$ , and  $\mathcal{M}$  represent collections of variables that determine the allocation of communication resources, the UV's positioning, and other mission-specific optimization variables, respectively. Then a high-level formulation is shown in (10),

$$\begin{aligned}
& \max_{\substack{\mathcal{R}, \mathcal{T}, \\ \mathcal{M}}} f_{\text{GLB}} \left( \underbrace{f_{\text{com}}(\mathcal{R}, \mathcal{T})}_{\text{Communication}}, \underbrace{f_{\mathcal{K}}(\mathcal{R}, \mathcal{T})}_{\text{Motion}}, \underbrace{f_m(\mathcal{R}, \mathcal{T}, \mathcal{M})}_{\text{Mission}} \right) \\
& \text{s.t. } \left. \begin{aligned} g_{\text{com}}(\mathcal{R}, \mathcal{T}) &\leq 0 \\ h_{\text{com}}(\mathcal{R}, \mathcal{T}) &= 0 \end{aligned} \right\} \text{Communication constraints,} \\
& \left. \begin{aligned} g_{\mathcal{K}}(\mathcal{T}, \mathcal{M}) &\leq 0 \\ h_{\mathcal{K}}(\mathcal{T}, \mathcal{M}) &= 0 \end{aligned} \right\} \text{Motion constraints,} \\
& \left. \begin{aligned} g_m(\mathcal{R}, \mathcal{T}, \mathcal{M}) &\leq 0 \\ h_m(\mathcal{R}, \mathcal{T}, \mathcal{M}) &= 0 \end{aligned} \right\} \text{Mission constraints,}
\end{aligned} \tag{10}$$

The objective may include communication, motion, and mission terms, combined as a weighted sum or in ratios. Not all terms will be present in all formulations, and furthermore, performance in any one of these areas may also be captured as constraints.

Communication constraints come in two types. First, there are communication resource constraints, which ensure that time, bandwidth, or spatial degrees of freedom are not over-allocated. Second, there are constraints tied to communication related metrics, including constraints on the minimal latency, SNR, or outage probability. Importantly, when a stochastic channel model is used, the constraints must be defined either in expectation or in some other probabilistic sense (*e.g.*, the constraint is met with some probability).

Kinematic constraints can also be subdivided into a few categories. First, there are kinematic constraints related to what movement is physically possible by the UV. Second, there may be safety constraints, such as a minimal flying height for a UV or a minimum distance between any two UVs. Finally, there may be constraints related to energy consumption for motion.

Finally, the mission-specific problem may introduce additional constraints or objectives. For example, if the UV is tasked with visiting a number of sites, we may additionally have a constraint that captures this requirement, or the objective may be to do so in the minimum amount of time, subject to, *e.g.*, energy and power constraints. These constraints may also ensure that mission objectives, such as sensing quality, meet certain thresholds.

### III. AN OPTIMIZATION TOOLKIT

In the previous section, we highlighted key modeling considerations and developed a high-level problem formulation which captured the many different optimization problems that arise when considering the operation of UVs in next-generation communication networks. However, it is difficult to develop a similarly generic solution method due to the wide variety of specific problem instances represented by the generic formulation. Rather, this section presents a number of important tools which find broad application, and these are classified into two categories. The first category comprises analytical methods, which generally produce interpretable solutions with well-defined optimality properties. The second category encompasses learning-based approaches, particularly in the context of stochastic dynamic programming with stochastic function approximation. These powerful methods are capable

of solving otherwise-intractable problems at the expense of interpretability of the solution.

#### A. Analytical Methods

Many optimization tools have been developed in fields such as computer science, operations research, robotics, and control, and set forth well-developed frameworks with clear understanding of optimality guarantees and computational complexity. Below, we discuss facets of UV operation that add complexity to problem formulations and highlight useful analytical tools for tackling them.

##### Handling Non-Convexity with Convex Approximation:

In many instances, meaningful measurement of performance requires consideration of nonlinear objectives or constraints (see Section II). For example, spatially non-convex channel gains or non-convex energy models, such as that of Eq. (8), create non-convexity in trajectory design problems.

For non-convex objective functions, *successive convex approximation* (SCA) provides an iterative method that guarantees a locally-optimal solution under permissive technical assumptions [95]. Specifically, for an objective function  $f(x)$ , the method finds a sequence of intermediate solutions  $\{x_k\}$  by solving a series of simpler problems,  $x_{k+1} = \arg \min_{x \in \mathcal{X}} \tilde{f}(x|x_k)$ , where  $\mathcal{X}$  is the solution space and  $\tilde{f}(x|x_k)$  is a strongly convex local approximator of  $f$ , possibly parameterized by the current iterate  $x_k$ . Newton's method is a common example of SCA, in which  $\tilde{f}(x|x_k)$  is the second-order Taylor-series expansion around the current iterate,  $x_k$ .

In some instances, coupling between optimization variables makes the problem difficult to solve. In such cases, *block coordinate descent* (BCD) simplifies the problem by partitioning the variables into non-overlapping subsets (blocks) and iteratively optimizing over a single block at a time. When the problem is convex in the block variables, standard convex optimization methods can be used. Otherwise, methods like SCA can be used to update a single block. Various optimality results hold under certain assumptions (see [96] and references therein). For UV problems, treating positioning and communication optimization variables as distinct blocks is often a useful way to decompose the problem.

##### Handling Combinatorics with Mixed Integer Programs:

Modeling concepts such as user-BS association or visit sequence among a set of WSN cluster heads introduces binary variables, giving these problems a combinatorial aspect. Mixed integer programs (MIPs), a broad class of mathematical programs characterized by the presence of both continuous- and integer-valued variables, can model these scenarios. While the general MIP is NP-hard, for special cases, extensive literature offers established methods for finding high-quality solutions.

Within the umbrella of MIPs, there are several important subclasses of problems characterized by the nature of the objective and constraint functions. In mixed-integer linear programs (MILPs), the objective and all constraint functions,  $f_i$ , are linear, whereas in a mixed-integer nonlinear program (MINLP) [97], the objective and/or constraints may be non-linear. Within the class of MINLPs, we further have a mixed-integer quadratic program (MIQP) [98] if the objective is

quadratic and the constraints are linear, and a mixed-integer second-order cone program (MISOCP) [99] if the objective is linear and all constraints are expressed as second-order cone constraints.

Branch-and-cut or branch-and-bound methods are common candidates for finding solutions of MIPs. These methods iteratively tighten upper or lower bounds on the optimal solution via relaxations of the integer variables or the non-linear constraints/objective. In the case of MILPs, the resulting relaxations are linear programs (LPs), which can be solved very quickly, so that optimal solutions may be found quite efficiently. Similarly, the convexity of the relaxations of MIQPs, MIQCPs and MISOCPs allow reasonably-sized problem instances to be solved optimally in a modest amount of time. A number of commercially available software packages are capable of solving these types of problems, including IBMs CPLEX [100], MOESK [101], and GUROBI [102]. For general MIPs, however, finding optimal solutions is quite challenging, and finding high-quality approximate solutions is often acceptable.

#### **Finding Visit Orders Using Vehicle Routing Problems:**

The vehicle routing problem (VRP) [103] refers to a general set of problems in which one or more vehicles must visit a number of locations in the shortest time possible. These problems may be posed as MIPs, with integer variables giving the sequence of site visits, and they include the classical traveling salesperson problem (TSP). Time window, energy, and vehicle-capacity constraints may be integrated into the problem as well.

The core modeling concept of the VRP is a weighted graph, with nodes representing locations to visit and edge weights giving the cost to move between sites. Works such as [104]–[106] provide an overview of the application of these types of problems for UV trajectory design. This framework can be made communication aware by incorporating communication aware metrics into the values of the edge weights.

#### **Solving Shortest-Path Problems with Graph-Theoretic Algorithms:**

Many problems of interest can be cast as shortest-path problems, particularly for scenarios where the UV is a consumer of network connectivity. For example, a cargo UAV may need to find the shortest path to its destination while respecting a QoS constraint or minimizing handoffs across BS. In these cases, the space may be discretized and modeled abstractly as a graph, and A\* [107] or similar graph-based algorithms can produce optimal solutions.

Randomized methods from the robotics community provide interesting alternatives. Rapidly-expanding random trees (RRT) find paths between given start and end positions by building out trees of partial solutions constructed from random samples of the UVs configuration space. These methods can quickly find feasible solutions, then iteratively improve, making them strong candidates for real-time applications, where available time for computation is unknown and short. Furthermore, they naturally handle obstacle avoidance as well as kinematic constraints, and a variant, RRT\*, provides asymptotic optimality guarantees [108].

**Sequential Decision-Making with Dynamic Programming and Optimal Control:** Both the VRP and shortest-path

problems are examples of more general sequential decision-making problems, for which dynamic programming (DP) and optimal control provide highly developed frameworks in discrete- and continuous-time settings, respectively. While some DP problems can be solved via analyzable and provably-optimal methods such as value iteration, many problems necessitate a learning-based approach. We discuss these methods further in the next subsection, but before doing so, we highlight the use of analytical optimal control.

When the dynamics of the problem are described with ordinary differential equations (ODEs), *e.g.*, when using a more developed kinematics model, optimal control [109] offers an established framework for planning the UV’s operation. Two key entities must be well-defined before posing the optimal control problem. First, an ODE in the form  $\dot{x}(t) = f(x(t), u(t))$  describe the evolution of the trajectory,  $x(t)$ , over time, with different control inputs,  $u(t)$ , producing different trajectories. Second, a cost functional

$$J(u(t)) = \int_{t=t_0}^{t_f} L(x(t), u(t)) dt + K(x(t_f), t_f)$$

provides a way to compare the quality of different inputs, where  $L$  and  $K$  are functions capturing the instantaneous and terminal costs/rewards, respectively.

As discussed in [110], optimal control can be conceptualized as DP in continuous time. Analogous to the Bellman equation, the Hamilton-Jacobi-Bellman equation characterizes the optimal control policy, and when the cost functional has special structure, more specific results may be applied. For example, if  $f(x(t), u(t))$  describes a linear time-invariant system and  $L(x, u)$  can be expressed in the quadratic form  $x^T Q x + u^T R u + 2x^T N u$  where  $Q$  and  $R$  are symmetric positive-definite matrices, and if there are no terminal costs, *i.e.*,  $K(x, u) = 0$ , we have a linear-quadratic regulator (LQR) problem, for which solutions may be found by solving an algebraic Riccati equation [111].

### *B. Learning-based Methods*

Trajectory design for UVs often requires optimizing system parameters, addressing long-term operation, and managing stochastic uncertainties. Stochasticity is inherent in this domain, particularly due to difficult-to-predict spatial and temporal channel variations brought about by shadowing and multipath effects [112]. Joint mobility of end-users and UVs further induces additional uncertainty. By taking a data-driven perspective and including stochastic function approximation, learning-based methods provide for the optimization of UV-assisted communication systems while ensuring robustness against uncertainty [113].

In this context, the Reinforcement Learning (RL) paradigm [114], [115] considers an agent that seeks to learn a policy for choosing actions by interacting with the environment. The agent’s goal is to maximize a notion of cumulative reward over time, more specifically, the expected sum of rewards that the agent collects over a horizon of interactions.

Mathematically, the RL problem can be formulated as a Markov Decision Process (MDP). An MDP is defined by a tuple  $(S, \mathcal{A}, \mathcal{P}, R, \gamma)$ , where:

- $\mathcal{S}$  is the set of states representing the environment's observable conditions.
- $\mathcal{A}$  is the set of actions that the agent can take to interact with the environment.
- $\mathcal{P}(\mathbf{s}_{k+1}|\mathbf{s}_k, \mathbf{a}_k)$  is the transition function, specifying the probability of transitioning to state  $\mathbf{s}_{k+1}$  given the current state  $\mathbf{s}_k$  and action  $\mathbf{a}_k$ .
- $R(\mathbf{s}_k, \mathbf{a}_k)$  is the reward function that provides a scalar feedback signal to the agent after taking action  $\mathbf{a}_k$  in state  $\mathbf{s}_k$ . An instantiation of the reward function at time step  $k$  can be denoted as  $r_k$ .
- $\gamma \in [0, 1]$  is the discount factor, representing the agent's preference for immediate rewards over future rewards.

The agent's goal is to learn an optimal policy  $\pi^*$  that maps states to actions, such that the expected cumulative reward, also known as the return, is maximized:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{\pi} \left[ \sum_{k=0}^{\infty} \gamma^k R(\mathbf{s}_k, \mathbf{a}_k \equiv \pi(\mathbf{s}_k)) \right]. \quad (11)$$

The state-action value function  $Q^{\pi}$  is defined as the expected discounted sum of rewards starting from a state-action pair and following the policy  $\pi$  thereafter, i.e.,

$$Q^{\pi}(\mathbf{s}, \mathbf{a}) = \mathbb{E}_{\pi} \left[ \sum_{k=0}^{\infty} \gamma^k R(\mathbf{s}_k, \pi(\mathbf{s}_k)) | (\mathbf{s}_0, \mathbf{a}_0) = (\mathbf{s}, \mathbf{a}) \right]. \quad (12)$$

The optimal value function  $Q^*(\mathbf{s}, \mathbf{a})$  is the fixed point of the Bellman backup operator [116]:

$$Q^*(\mathbf{s}, \mathbf{a}) = \mathbb{E}_{\mathbf{s}' \sim \mathcal{P}} \left[ R(\mathbf{s}, \mathbf{a}) + \gamma \max_{\mathbf{a}'} Q(\mathbf{s}', \mathbf{a}') \right]. \quad (13)$$

In Deep RL (DRL) [117], the optimal state-action value function is often parameterized as a neural network with parameters  $\mathbf{w}$ , denoted as  $Q_{\mathbf{w}}(\mathbf{s}, \mathbf{a})$ , also known as the *critic* or *value network*. Approximating the value function involves using batches of transitions, typically stored in an Experience Replay  $D_{\text{ER}}$ , and applying the following gradient descent update rule on the critic:

$$\mathbf{w} \rightarrow \mathbf{w} + \eta \mathbb{E}_{(\mathbf{s}, \mathbf{a}, \mathbf{s}', r) \sim D_{\text{ER}}} \left[ \left( Q_{\mathbf{w}}^*(\mathbf{s}, \mathbf{a}) - Q_{\mathbf{w}}(\mathbf{s}, \mathbf{a}) \right) * \nabla_{\mathbf{w}} Q_{\mathbf{w}}(\mathbf{s}, \mathbf{a}) \right], \quad (14)$$

where  $*$  is the standard multiplication operator and the scalar parameter  $\eta$  denotes the learning rate.

In the context of MDPs with discrete action spaces, achieving an accurate estimate of the value function is deemed sufficient for optimal control [118]. This arises from the fact that, at each state  $\mathbf{s}_t$ , the agent can straightforwardly select the action  $\mathbf{a}_t$  that maximizes the value estimator.

Discrete action spaces are prevalent in UV problems, often resulting from the discretization of the UV's operational space, thus inducing a discrete action space MDP. Moreover, in various scenarios, the action space is inherently discrete. For instance, in situations where decisions like the UV-BS or the UV-end user associations are made at every time step, the action space inherently takes on a discrete nature.

However, when dealing with continuous action spaces, the maximization with respect to the action requires a different approach. In such a scenario, action selection is achieved by parametrizing the policy with a neural network with parameter

vector  $\phi$ ,  $\pi_{\phi}(\mathbf{a}_k|\mathbf{s}_k)$ . The parameter vector  $\phi$  is adjusted with stochastic gradient descent to choose the action at each state that maximizes the estimation of the critic [119]:

$$\nabla_{\phi} J(\phi) \approx \mathbb{E}_{\mathbf{s}_k \sim \mathcal{P}} \left[ \nabla_{\mathbf{w}} Q_{\mathbf{w}}(\mathbf{s}_k, \pi_{\phi}(\mathbf{s}_k)) \nabla_{\phi} \pi_{\phi}(\mathbf{s}_k) \right]. \quad (15)$$

In UV-assisted wireless communications, decision variables like energy expenditure, motion space, and bandwidth allocation inherently take values from a continuous set. Discretizing the solution set is conventionally non-scalable due to the curse of dimensionality [120]. Thus, a more effective strategy is to use continuous action space DRL approaches for addressing these complex problems.

Various versions of the policy gradient algorithm adapt gradient estimation for different purposes. Proximal Policy Optimization (PPO) [121] employs a clipped surrogate objective to maintain stability during policy updates, preventing large changes when value estimates are noisy. Similarly, Trust Region Policy Optimization (TRPO) [122] confines policy updates within a trust region of the parameter space to ensure smooth transitions in learning steps.

Certain methods avoid using a function approximator to parameterize the value function and concentrate solely on parametrizing the policy. These approaches replace the parametrization of the critic with Monte Carlo estimates of the value function, fitting within the Baseline Policy Gradient (BPG) framework. Here, the policy is often presented as a learnable distribution over the action space, and the baseline policy gradient is usually formulated as follows:

$$\nabla_{\phi} J(\phi) \approx \mathbb{E}_{\mathbf{s}_k \sim \mathcal{P}} \left[ \hat{R}_k \nabla_{\phi} \log \pi_{\phi}(\mathbf{s}_k) \right], \quad (16)$$

where  $\hat{R}_k$  is the Monte Carlo estimate of the expected sum of rewards corresponding to the state-action pair  $(\mathbf{s}_k, \mathbf{a}_k)$ . Since Monte Carlo estimates usually exhibit high variance, the baseline policy gradient methods are considered unstable.

1) *On-policy vs Off-policy*: A fundamental distinction exists between *On-policy* and *Off-policy* DRL methods. On-policy methods optimize the current policy based on its generated data, discarding experiences from previous policies, which can slow learning. In contrast, Off-policy methods decouple data collection from optimization, using Experience Replay [117] to store past experiences for more sample-efficient updates. While On-policy methods like TRPO [122] and PPO [123] prioritize stability, Off-policy methods (e.g., Deep Q-Learning (DQL) [117] and Soft Actor-Critic (SAC) [124]) often demonstrate greater sample efficiency and flexibility in learning from diverse experiences.

2) *Model-based vs Model-free*: The dynamics of the MDP are encapsulated in the probability distribution over subsequent states and rewards given the current state-action pair  $\mathcal{P}(\mathbf{s}_{k+1}, r_k|\mathbf{s}_k, \mathbf{a}_k)$ . The presence or absence of assumptions concerning these dynamics delineates the distinction between *Model-based* and *Model-free* DRL approaches.

In Model-based DRL, the agent has an internal representation of transition dynamics, either assumed beforehand or estimated through data-driven techniques [125], [126]. With this insight, the agent can simulate trajectories to determine the optimal policy. In Model-free approaches, there is no internal representation of the underlying dynamics. Instead, the agent

directly learns the value function and optimal policy through an iterative trial and error process.

Exploring the intersection of UV-assisted wireless communications and DRL reveals the distinction between Model-based and Model-free approaches, especially regarding the underlying channel statistics. In other words, to comprehend the MDP dynamics, the UV agent requires specific assumptions about the statistical evolution of communication channels over time and space [127], [128].

3) *DRL Algorithms*: DRL is built on a core theoretical framework, but there are many specific algorithms which handle the various cases discussed above. Table III summarizes the characteristics of various Model-free DRL algorithms. Applications of these methods to optimize UV operation in communication systems are found in Section IV.

TABLE III  
THE CHARACTERISTICS OF VARIOUS MODEL-FREE DRL APPROACHES.

Algorithms	On-Policy	Off-Policy	Critic	Policy
DQL [117]		✓	✓	
Dueling DQL [129]		✓	✓	
SAC [124]		✓	✓	✓
TD3 [130]		✓	✓	✓
PPO [121]	✓		✓	✓
TRPO [122]	✓		✓	✓
DDPG [131]		✓	✓	✓
BPG [114], [115]	✓			✓
Zeroth-order DPG [132]	✓			✓

#### IV. REPRESENTATIVE PROBLEM FORMULATIONS

As made apparent in previous sections, the set of possible problem formulations that arise when optimizing the operation of UVs in next-generation networks is incredibly large and diverse. This diversity stems not only from the myriad roles and applications of UVs in the network, as discussed in Section I, but also from the range of modeling choices, technical assumptions, and optimization tools that may apply to a single problem, as discussed in Sections II and III.

Despite this variety, work in this area naturally clusters into a set of core problems, which we illustrate next with a series of representative problem formulations. These are presented with a level of abstraction so as to be agnostic to, *e.g.*, UV kinematics or channel modeling, while highlighting the fundamental question posed in each. We first present analytic optimization formulations (see Section III), then present a selection of these in the context of reinforcement learning. Throughout, we discuss the importance of each problem and give examples from the literature.

##### A. Unmanned Vehicles as Providers

In this subsection, we present three broad categories of problems in which UVs act as providers. In the first, UVs provide connectivity to the core network for multiple users, possibly in environments which include terrestrial BS (TBS) as well. In the second, the UV uses ISAC to perform a sensing task while simultaneously communicating. In the third, UVs work cooperatively to extend connectivity by relaying either with multihop communication or distributed beamforming.

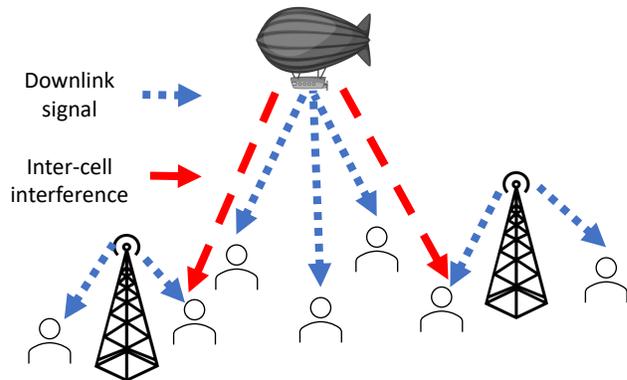


Fig. 6. A UV acting as a base station must support concurrent users while mitigating inter-cell interference.

##### 1) Providing Multiple Access with Resource Allocation:

When UVs act as base stations BSs (UV-BS), either as a cost-effective solution to closing the digital divide or as dynamic assets to relieve temporary high-traffic loads in urban areas, a core problem is resource allocation: What is the best way to distribute a UV-BS's finite wireless resources (power, time, bandwidth, and spatial degrees of freedom) across multiple concurrent users?

Resource allocation over multiple simultaneous links for UV-BS differs from existing results on multi-access for TBS in two important ways. First, the mobility of UV-BSs creates an additional degree of freedom for planning. Second, the elevated position of aerial UV-BSs results in a high probability of a strong LoS channels to unintended receivers, so that mitigating inter-cell interference takes on greater importance. A high-level depiction of the problem setting is shown in Figure 6.

To pose a general problem, let  $\mathcal{U}$  denote the set of users,  $C_i, i \in \mathcal{U}$  indicate a per-user communication metric (Section II), and  $f_{\text{comm}}$  here be an abstract aggregation operator which takes as input each user's link-level performance and produces a single, system-level metric. Let  $\mathcal{B}_{\text{UV}}$  denote the set of UV-BS,  $\mathcal{B}_{\text{TBS}}$  the set of TBS, with  $\mathcal{B} = \mathcal{B}_{\text{UV}} \cup \mathcal{B}_{\text{TBS}}$ . The problem consists of finding the binary BS-to-user associations,  $\{a_{bi}\}_{(bi) \in \mathcal{B} \times \mathcal{U}}$ , the locations of the UV BS,  $\{q_b\}_{b \in \mathcal{B}_{\text{UV}}}$ , and the resources allocated from each BS to each user,  $\{\mathcal{R}_{bi}\}_{(bi) \in \mathcal{B} \times \mathcal{U}}$ , which maximize the aggregate metric of system performance subject to positioning, communication resource, and QoS constraints. The problem is given below:

$$\max_{\substack{\{a_{bi}\}, \\ \{q_b\}, \{\mathcal{R}_{bi}\}}} f_{\text{com}} \left( \{C_i(\{a_{bi}\}, \{\mathcal{R}_{bi}\}, \mathcal{H}_i(\{q_b\}))\}_{i \in \mathcal{U}} \right) \quad (17a)$$

s.t.

$$g_{\text{com}}(\{\mathcal{R}_{bi}\}) \leq 0, \forall b \in \mathcal{B}, \quad (17b)$$

$$\tilde{C}_i(\{a_{bi}\}, \{\mathcal{R}_{bi}\}, \mathcal{H}_i(\{q_b\})) \geq \tilde{C}_{i, \text{req}}, \forall i \in \mathcal{U}, \quad (17c)$$

$$\sum_{b \in \mathcal{B}} a_{bi} = 1, \forall i \in \mathcal{U}, \quad (17d)$$

$$a_{bi} \in \{0, 1\}, \forall b \in \mathcal{B}, i \in \mathcal{U}, \quad (17e)$$

$$q_b \in \mathcal{Q}_b, \forall b \in \mathcal{B} \quad (17f)$$

where, for brevity, we have written the association and resource allocation variables as  $\{z_{bi}\}$  rather than  $\{\{z_{bi}\}\}_{(bi) \in \mathcal{B} \times \mathcal{U}}$  for  $z \in \{a, \mathcal{R}\}$ , and  $\mathcal{H}_i(\{q_b\})$  represents

TABLE IV  
SAMPLE PAPERS ADDRESSING THE OPTIMIZATION OF UVs IN THE ROLE OF PROVIDERS.

UV Mission	Ref.	Objective	Tools	Notes
UV as BS	[133]	Worst-case SINR	MIP, SCA	OFDMA, sub-channel assignment, HAPS and TBSs.
	[134]	Sum-rate	SCA	PD-NOMA, maritime environment.
	[135]	Number of users	MINLP	OMA, 3D placement.
	[136]	Sum-rate	SCA, Semi-definite relaxation	MIMO, full-duplex communication.
	[137]	User-avg. latency	Optimal Transport	OFDMA, non-terrestrial network (NTN).
	[138]	Worst-case Throughput	BCD/SCA	TDMA, trajectory planning.
	[139]	Total communication rate	RL	SIMO, Q-learning for motion control in order to maximize communication rate.
	[140]	Throughput	DRL	MIMO, DPG for weighted throughput maximization subject to fairness.
	[141]	Energy consumption	DRL	Motion Control of UV base stations in order to minimize energy expenditure subject to a constraint of perpetual connectivity.
	[142]	Task Completion	TD-learning	TD-learning with tile coding for motion control of UV base stations.
	[143]	Sum-rate	DRL	UV continuous motion control for collective sum-rate maximization.
	[144]	Throughput	DRL	MIMO, SARSA and offline RL for throughput maximization under fading.
ISAC	[145]	Generic sensing and communication	Calculus of variations	Sensing performance measured with sensing SNR.
	[146]	Sum-rate	MIP, BCD, SCA	MIMO, SDMA for sensing and communication.
Cooperative Relay	[75], [147]	SINR	Distributed stochastic optimization	Distributed beamforming. A queue of previous positions and channels is used to solve a stochastic surrogate for channel prediction.
	[148]	Bit-error probability, motion energy	KKT conditions	Multihop robotic relay, uncertain channel environment with obstacles, Gaussian-process based channel prediction.
	[149]	Abstract task completion	SOCP, control barrier functions, RRT*	Multihop relay. Extensive real-world experimentation.
	[150], [151]	Motion and communication energy	Multiple-choice knapsack problems	Distributed robotic beamforming. Accounts for uncertainty in shadowing and multipath fading via stochastic channel prediction.
	[152], [153]	SINR	DRL	Distributed beamforming relay. Fourier features as preprocessing to estimate the value function depending on spatio-temporally correlated channels.
	[154]	Secrecy rate	Stochastic programming	Distributed beamforming relay. Distributed stochastic programming for secrecy rate maximization under cooperation.
	[155]	Throughput	DRL	Single UV relay. Motion control for average system throughput maximization.
	[156]	Transmit power	BCD	BCD to optimize UV positioning in multihop relaying.
	[157]	Maintaining connectivity	Particle Swarm Optimization	Particle Swarm Optimization for UV placement for multihop relaying.
	[158]	Min communication rate	BCD, Gibbs sampling	3D UV placement and power and bandwidth allocation for multihop relaying.
[159]	Energy	Multi-agent DRL	Single UV relay. Multi-agent RL surpasses hybrid precoding for energy minimization.	

the set of abstract channels from all BS to user  $i$  (see Section II). The link-level metric  $\mathcal{C}_i$  is a function of BS-to-user associations as well as resource allocation and the communication channels as determined by the BS positions, and (17c) is a QoS constraint which requires that a link-level metric  $\tilde{\mathcal{C}}_i$  exceeds a value  $\tilde{\mathcal{C}}_{i,\text{req}}$  for each user. Constraint (17b) ensures communication resources are not over-allocated, and (17f) captures positioning constraints such as obstacles, no-fly zones or fixed flying altitudes.

There are several important challenges raised in this generic problem statement. The binary user association variables introduce a combinatorial element to the problem, which is naturally expressed as a MIP. Non-convexity enters through inter-cell interference, highlighted by the dependence of  $\mathcal{C}_u$  and  $\tilde{\mathcal{C}}_u$  on *all* resource allocation variables, as well as through non-convex spatial variations in the communication channels, which enter the problem through  $q_b$ .

Many problems may be formulated through specific choices of the link-level metric, aggregation operator, resources allocation constraints, and positioning constraints. We next review a few interesting examples.

**Providing Multiple Access with Given Position:** In some cases, such as when dealing with quasi-stationary HAPSs, the positioning of the UV is assumed fixed, so that optimization only occurs over the communication resource variables. For example, in [133], a system consisting of multiple TBSs and a single HAPS serve a set of ground users using OFDMA. The authors formulate a problem which seeks to maximize the minimum SNR over all users by optimally selecting a sub-carrier frequency channel and power allocation for each user. An iterative procedure based on successive convex approximation (SCA) repeatedly solves a MIP to find solutions to the problem which enhance the worst-case SNR among the users.

A multi-UV variation appears in [134], in which a team of UAV-BSs, TBSs, and a satellite jointly provide communication to several maritime users using non-orthogonal multi-access (NOMA). The transmit powers of the UAV-BSs and TBSs are jointly optimized to maximize the sum-rate under the constraint that interference does not significantly degrade communication for the users connected via satellite. To solve the problem, an iterative SCA method is proposed, and results show that the NOMA multi-access scheme greatly improve the sum-rate compared to OMA schemes in this maritime scenario.

**Multiple Access with Optimized Positioning:** In many cases, a UV-BS deployed to service an area will operate from or near a fixed point. In [135], the authors examine the 3D-placement problem for a UAV which is deployed to improve the network capacity. The UAV services all users for which the channel between the user and the UAV achieves some threshold value, and the 3D placement of the UAV must be chosen to maximize the total number of users serviced by the UAV. The problem is formulated as a mixed-integer nonlinear program (MINLP) which is solved using an interior-point method and a bisection search.

In addition to UAV placement, [136] optimizes beamforming weights and uplink-user transmit power in a scenario where a full-duplex UAV-BS simultaneously handles uplink and downlink traffic from distinct sets of users with LoS/NLoS

mmWave channels. The challenging non-convex problem is solved by first optimizing the UAV's location using SCA, then finding the UAV-BS beamforming weights and uplink user transmit powers by relaxing a rank-1 matrix constraint. It is shown that omitting optimization of the UAV-BS's location reduces sum-rate by nearly 1 b/s/Hz, and that greater performance loss occurs when using only half-duplex.

The work in [137] develops a fully aerial network, with designated UAV-BSs providing connectivity to UAV end users using OFDMA. The authors pose the problem of partitioning the coverage region among the UAV-BS to achieve the minimal average communication latency, and the problem is solved using optimal transport [160].

**Multiple Access with Trajectory Optimization:** When optimizing over time, additional kinematic and energy constraints may enter the problem. We express these generically as

$$\left. \begin{aligned} g_{\mathcal{K}}(\{q[k]\}_{k=1}^K) &\leq 0 \\ h_{\mathcal{K}}(\{q[k]\}_{k=1}^K) &= 0 \end{aligned} \right\} \text{Motion constraints} \quad (18)$$

where  $k$  and  $K$  are the discrete-time index and horizon, respectively. For specific examples of such constraints, see Section II. For brevity, in the sequel we denote  $\{q[k]\}_{k=1}^K$  with  $\{q[k]\}$ , unless clarity demands otherwise.

The authors of [138] consider multiple UAV-BSs moving from given initial locations to given final locations while providing connectivity to a set of end users using TDMA. A problem is formulated which maximizes the minimal time-averaged throughput among end users by jointly optimizing trajectories, transmit powers, and user associations. To solve the problem, block coordinate descent (BCD) is used in conjunction with SCA to find solutions. Numerical examples show that utilizing mobility significantly increases the minimal throughput compared to the placement of static BSs.

**Takeaways:** Thus far, we have looked at resource allocation and positioning for UV-BSs supporting multiple users. These problems are characterized by binary user-BS association variables and a need to mitigate inter-cell interference. Several examples have illustrated the performance gains achievable by deploying UVs and exploiting their mobility, and SCA has proved to be a useful tool for solving these problems. We next consider integrated sensing and communication (ISAC), in which shared time, frequency, and power resources are allocated not only among different communication users but also more broadly between sensing and communication tasks.

2) **Integrated Sensing and Communication (ISAC):** Recently, there has been growing interest in integrating communication and sensing, particularly as communication systems begin to employ higher frequencies which are more amenable to use in radar-like ways. Fully integrating these two applications requires combining hardware, signal processing pipelines, and waveforms, as well as understanding how to allocate limited resources between these two tasks. A thorough survey of the area may be found in [14], [161], [162].

Many advantages of UVs for communication also extend to sensing-related tasks. Mobility allows for repositioning to find better sensing locations, while the high probability of LoS paths for UAVs/HAPSs avoids occlusion in the sensing

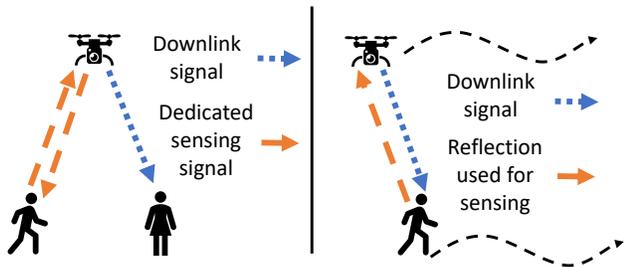


Fig. 7. ISAC may use either orthogonal sensing and communication signals, or a single signal may be used for both applications.

domain. [163] and [164] provide more discussion of ISAC specifically in the context of UAVs.

To pose a general problem, we consider two sets of resource-allocation variables,  $\mathcal{R}_C$  and  $\mathcal{R}_S$ , representing the resources allocated to communication and sensing, respectively. While current sensing standards treat these resources as orthogonal, as in 802.11bf which uses time-division to isolate sensing and communication signals [165], finding dual-function waveforms that achieve sensing and communication simultaneously is an area of active research. For a single UV, a general problem formulation is given below:

$$\max_{\substack{\{q[k]\}, \{\mathcal{R}_C[k]\}, \\ \{\mathcal{R}_S[k]\}}} f_{\text{GLB}}(f_{\text{com}}, f_{\text{sense}}) \quad (19a)$$

$$\text{s.t.} \quad (18) \text{ (Motion constraints),}$$

$$g_{\text{com}}(\mathcal{R}_C[k], \mathcal{R}_S[k]) \leq 0, \forall K. \quad (19b)$$

The objective in (19a) captures the trade-off between communication performance,  $f_{\text{co}}$ , and sensing performance,  $f_{\text{sense}}$ , and both are functions of all optimization variables. Either the sensing or communication performance could alternatively be treated as a constraint, and Constraint (19b) captures the fact that resources are shared across sensing and communication applications. We next provide two interesting examples.

In [145], the authors study the optimal trajectories for a UAV tasked with sensing and communicating with a mobile ground node. They show that the sensing and communication quality at different points can be viewed as an artificial potential field, so that maximizing an abstract communication and sensing service can be found by determining the shape of a catenary [166]. Assuming path loss-only communication models, analysis shows the global optimality, existence, and uniqueness of the given solution.

The sensing performance may be taken as a constraint rather than part of the objective. For example, the work in [146] investigates a multi-antenna UAV which communicates with multiple users using TDMA while also periodically sensing a number of targets. To optimize the total communication rate, communication user and sensing targets are scheduled along with corresponding beamforming weights, subject to the constraint that when sensing is scheduled, the sensing SNR (distinct from the communication SNR) must exceed a given threshold. The problem is solved using a general BCD approach, and results show a trade-off between achievable communication performance and sensing frequency.

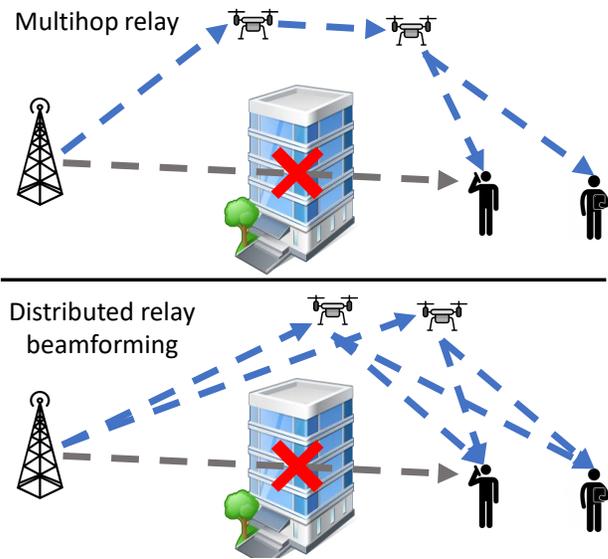


Fig. 8. UV-enabled distributed relay beamforming can extend existing terrestrial networks.

**Takeaways:** As with communication, UVs can use mobility to improve sensing in ISAC scenarios. However, they must balance resource allocation and trajectory planning between communication and sensing objectives, which may either be competing or complementary. Furthermore, ISAC is an emerging paradigm, and as the technology matures, the place of UVs in enabling these applications will continue to develop.

**3) Relaying and Cooperative Communication:** In many scenarios, UVs can work cooperatively to enhance communication in the network. A key example is the use of UVs as relays, either by creating a chain of communication links to extend connectivity via multihop relay or coordinating transmission to achieve distributed beamforming.

A generic setting is the following:  $M$  sources aim to communicate with  $M$  destination nodes, but direct communication is not possible due to obstruction of a LoS path or significant distance. Facilitating this communication are  $N$  UV relays which will move to locations  $q_n$ ,  $n \in \{1, \dots, N\}$ . The communication performance for each source-destination pair (SDP) is given by  $C_m$ ,  $m \in \{1, \dots, M\}$ , which represents metrics such as data rate or bit error rate (BER). For each SDP, let  $S_m$  be an abstract object representing the routing path from source to destination (e.g., a sequence of UVs), and let  $\mathcal{R}_{n,m}$  denote the communication resources, allocated by UV  $n$  to SDP  $m$ , with  $\mathcal{R}_m = \{\mathcal{R}_{n,m} | n = 1, \dots, N\}$ . Then, a high-level problem can be formulated as below:

$$\max_{\substack{\{q_i\}_{i=1}^N, \{\mathcal{R}_m\}_{m=1}^M, \\ \{S_m\}_{m=1}^M}} \sum_{m=1}^M C_m(\mathcal{R}_m, S_m, \{q_i\}) \quad (20a)$$

$$\text{s.t.} \quad \{q_i\}_{i=1}^N \in \mathcal{Q}, \quad (20b)$$

$$g_{\text{com}}(\{\mathcal{R}_m\}) \leq 0, \quad (20c)$$

$$g_r(\{S_m\}) \leq 0. \quad (20d)$$

Constraint (20b) ensures that the placement of the UVs is valid, e.g., enforcing minimum distances between UVs, while

Constraint (20c) ensures that communication resources are not over allocated. Constraint (20d) requires that communication routing is valid, e.g., there are no cycles. For clarity of exposition, we have written the SDP-level communication performance,  $C_m$ , as a function of only the resources allocated to SDP  $m$ , but more generally, e.g., when communication for the different SDPs does not occur over orthogonal channels,  $C_m$  may be a function of resources allocated to other SDPs as well. We next look at examples of these types of problems from the literature, focusing first on multihop routing, then proceeding to distributed beamforming approaches.

**Multihop mobile relaying:** Multihop relaying [167] (see Fig. 8 (top)), departs from the conventional point-to-point communication model by introducing a chain of data transmission through intermediary nodes. Such approaches can extend coverage and hedge against NLoS conditions. UVs can effectively serve as relays in multihop scenarios, as their inherent mobility and adaptability prove beneficial in responding to evolving communication conditions, such as variations in communication channels and the mobility of end-users.

As an example of such a scenario, [148] considers multiple UVs moving from initial positions to find locations which minimize the end-to-end bit error rate (BER) for a multihop relay scenario in the presences of obstacles and stochastic communication channels. For certain conditions, the problem is shown to be concave, and the KKT conditions can be used to produce solutions. [149] considers an ad hoc network of robots which must support multiple source-destination communication flows while performing an abstract task. For tractability, the problem is decomposed, with communication routing found using a second-order cone program (SOCP) [168], local control laws derived using barrier functions [169], and RRT\* [108] providing high-level trajectories.

In [156], the authors consider a problem such as the one described above. They propose a Block Coordinate Descent (BCD) [170] method to find the positions of the UVs which minimize the total transmit power under a coverage constraint. The work of [157] examines a similar UV multihop scenario and proposes a particle swarm optimization approach to design the UV positions for maintaining connectivity. Finally, the authors of [158] examine a scenario with multiple source-destination pairs where communication is facilitated by multiple UV relays in a multihop setting. They propose an alternating optimization scheme that combines BCD with Gibbs-sampling [171] to jointly optimize the UV positions as well as the bandwidth and power allocation, with the objective of maximizing the minimum communication rate among source-destination pairs.

**Distributed relay beamforming:** Distributed relay beamforming [172], akin to multihop relaying, involves source-destination pairs seeking information exchange. Relays between these pairs receive signals simultaneously and transmit them collectively to destinations in synchronous fashion. Despite its classification as a multihop relaying case with two hops, it is often studied independently due to nuanced domain-specific considerations. These nuances involve cooperative relay decision-making to form a unified beam at the destination.

In recent years, beamforming in the context of UVs has

gained significant attention. For example, [150], [151] focus on minimizing both motion and communication energy while jointly optimizing the positions of multiple UVs and their corresponding beam weights for distributed robotic beamforming. The paper considers an uncertain channel environment and shows how to incorporate channel prediction using a Gaussian-process-based prediction model. Near-optimal solutions are then found through a series of multiple-choice knapsack problems [173], which can be solved efficiently despite being NP-hard in the general case.

The study in [174] explores scenarios where multiple relays engage in distributed beamforming for a single source-destination pair, incorporating quantized feedback. The authors recommend using the Generalized Lloyd Algorithm [175] to design the quantizer for specified beamforming weights. In a related domain, [176] delves into employing multiple relays for distributed relay beamforming in a two-transceiver system. Two scenarios are considered: optimizing relay beamforming weights to minimize total transmit power while meeting receiver SNR constraints, and maximizing the lowest SNR between two receivers under a total transmit power constraint. Proven unique solutions and proposed iterative schemes showcase linear complexity per iteration. In cooperative jamming, [177] focuses on UV motion control and noise nulling to maximize collective secrecy rates. [154] optimizes relay positioning for acceptable receiver-side QoS and maximizes ergodic secrecy rates with an eavesdropper.

The works [75], [147] consider scenarios with one source-destination pair and multiple UV relays, aiming to select UV positions and beamforming weights to maximize a utility metric. Assuming underlying channels evolve as Gaussian processes with correlations across time and space, the authors propose a distributed stochastic optimization method. This approach predicts channels for candidate UV positions based on historical data, achieving closed-form expressions for optimal beamforming weights once positions are determined.

**Takeaways:** The use of UVs as mobile relay nodes in relay networks introduces the element of dynamic positioning, an additional controllable parameter for optimizing various aspects of system performance. When multiple UVs provide this relay service together, their operation must be optimized jointly. The cooperative paradigm, whether the team of UVs acts as a single relay chain or employs distributed beamforming, greatly influences optimal positioning and resource allocation.

## B. Unmanned Vehicles as Consumers

Supporting connectivity for UVs is an expected feature of 6G networks, with related standards actively under development [62]. In this section, we present several fundamental problem formulations in which UVs consume the services provided by 6G systems, beginning with the problem of maintaining connectivity, then moving towards more complex problems of wireless data-offload scheduling with data causality constraints in the context of IoT and WSN data collection.

1) **Maintaining Connectivity while in Motion:** Consider a UV which must move between two points while maintaining

TABLE V  
SAMPLE PAPERS ADDRESSING THE OPTIMIZATION OF UVs IN THE ROLE OF CONSUMERS.

UV Mission	Ref.	Objective	Tools	Notes
Maintaining Connectivity	[178]	Motion and communication energy	RRT*	Obstacle avoidance, uncertain channel environment.
	[179]	Mission time	Graph theory, Convex optimization	UAV associates with nearest BS.
	[180]	Handovers	Dynamic programming	Cargo UAV application.
Reconnecting to the Network	[181]	Distance to connectivity	Game theory, Graph theory, Dynamic programming	Probability of success modeled as independent across discretized space. Shown to be NP-hard.
	[182]	Distance to connectivity	Ornstein-Uhlenbeck process	Accounts for log-normal shadowing and generic multipath fading.
Upload Scheduling	[183]	Motion and communication energy	KKT conditions	Velocity and transmission rate optimized over a given trajectory.
	[184], [185]	Motion and communication energy	Optimal control	Abstract Bolza optimal control problem.
Site Visiting	[186]	Sum-rate and completion time	DRL	Deep Q-learning for motion control to optimize a convex combination of sum-rate and task completion time.
	[187]	Avg. Wait Time	Polling systems, MISOCP, stochastic planning	Data collection and relay. Optimal robotic routing locations/policies. Long-run analysis of stationary regime
	[188]	Avg. Wait Time	Polling systems, DRL	Data collection and relay. Dynamic visit sequence. Handles long-run average cost criterion.
	[189]	Data collection and wireless energy transfer	Multi-agent DRL	Split the UVs into two groups (data harvesting and energy transfer) and employ multi-agent DRL since the policies are correlated.
	[190]	Data rate	DRL for partial observability	Buffer overflow can cause unsuccessful transmission that induces partial observability.
	[191]	Motion and communication energy	MILP	Sensing mission, collects set amount of data from each site.
	[192]	Motion and communication energy	Monte Carlo tree search	Trajectory must pass within a certain distance of each site.
	[193]	Device energy	Sequential quadratic programming	Energy reduction for IoT device communication
	[194]	Number of devices served	MINLP, SCA, Branch, reduce, & bound	Shown to be NP-hard.
	[195]	UV energy	Vehicle routing problem	Capacitated vehicle routing problem with time windows.
	[196]	Motion and communication energy	MILP	Sites visited on Hamiltonian path.
	[197]	Peak/Avg. AoI	Ant colony optimization	Sensor-node-to-UAV wireless communication, UV moves to BS before offloading (data muling).
[198]	Avg. AoI	BCD, SCA, Convex optimization	Considers communication over both WSN-to-UAV and UAV-to-BS links.	

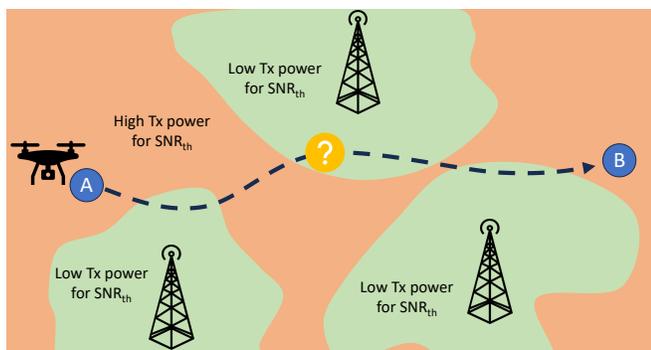


Fig. 9. Maintaining connectivity: A fundamental problem for UVs is how to efficiently move between two locations while maintaining connectivity.

connectivity with a remote server in order to, *e.g.*, receive control and non-payload communication (CNPC) [199] or upload

sensed data in real time. These problems can be frequently posed as shortest-path problems, with distance metrics that account for transmission and motion power. As a baseline example, consider a UV moving from a given initial position,  $q_0$ , to a final position,  $q_f$ , while maintaining the SNR between the UV and a remote BS above a minimum threshold,  $\text{SNR}_{\text{th}}$ , as illustrated in Figure 9. This baseline formulated is

$$\min_{\substack{\{q[k]\}_{k=1}^K, \\ \{\mathcal{R}[k]\}_{k=1}^K}} \lambda_1 \mathcal{E}_{\mathcal{K}}(\{q[k]\}) + \lambda_2 \mathcal{E}_{\text{com}}(\{\mathcal{R}[k]\}) + \lambda_3 \Delta t K \quad (21a)$$

s.t. (18) Motion constraints,

$$g_{\text{com}}(\mathcal{R}[k]) \leq 0, \forall k, \quad (21b)$$

$$\text{SNR}(q[k], \mathcal{R}[k]) \geq \text{SNR}_{\text{th}} \forall k, \quad (21c)$$

$$q[0] = q_0, \quad (21d)$$

$$q[K] = q_f \quad (21e)$$

where,  $\Delta t$  is the length of each time step,  $\mathcal{E}_{\mathcal{K}}(\{q[k]\})$  is the motion energy required to follow trajectory  $\{q[k]\}$ , as determined by the UV's energy model (see Section II),  $\mathcal{E}_{\text{com}}$  is the communication energy,  $\Delta t K$  is the mission completion time, and  $\lambda_i, i \in \{1, 2, 3\}$  are scalar weights. Thus, the objective consists of a weighted sum of motion energy, communication energy, and mission completion time, with the latter particularly important in, *e.g.*, drone delivery problems. We can formulate many related problems by moving performance metrics between constraints and objectives. For example, the SNR constraint can be relaxed by including a penalty term in the objective for the distance or time traveled with  $\text{SNR} < \text{SNR}_{\text{th}}$ , and the mission completion time can be recast as a constraint on the maximum allowed time.

A version of this problem is studied in [178], in which a UV moves through a complex environment with obstacles while maintaining connectivity to the core network via any one of several TBSs. Both fixed and variable transmit power problems are considered, and due to the additive nature of the objective function, the problem can be solved using a modified version of RRT\*. The resulting trajectories conserve communication energy by avoiding areas of poor channel quality. In a similar setting, [179] studies the problem of a UAV flying between an initial and final location in minimal time while maintaining connectivity with a set of TBSs. The authors find high-quality approximate solutions using tools from graph theory and convex optimization.

When the UV may choose from among multiple BS for connectivity, it should avoid excessive handoffs, which are expensive from the standpoint of communication resources. This is accounted for in [180], which studies the problem of a cargo UAV flying at a fixed altitude between two points which seeks to minimize a linear combination of energy consumption and number of handoffs while constrained to be disconnected no more than a given percentage of the path. The path planning problem is solved using a dynamic programming approach. As the flight altitude increases, the energy consumption increases and handoffs decrease, while stricter disconnectivity constraints also increase energy consumption.

**Takeaways:** Maintaining connectivity can be viewed as a shortest-path problem with an appropriate choice of distance metric. Consequently, graph theory and DP provide good approaches to solving these problems. The resulting paths improve communication by seeking out areas of high channel quality, possibly at the expense of additional motion energy.

2) **Minimizing Distance to Connectivity with Stochastic Channel Models:** When a UV loses connectivity, either while fulfilling a mission objective or due to the difficult-to-predict variations in the communication channel, it may need to quickly return to an area where communication is possible. Here, a problem naturally arises: how can the UV minimize the time it takes to find a connected point?

Consider a UV initially at a position  $q_0$  where the channel quality does not permit communication with a distant BS. Without loss of generality, assume the UV moves at a constant velocity,  $v$ , and transmits with a fixed transmit power,  $P$ . The channel is not known, and at each point, the condition (21c) is met with some probability tied to the spatially-correlated

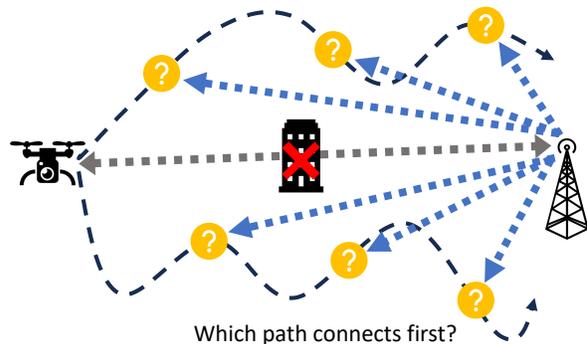


Fig. 10. Finding a connected spot: For a UV that is initially disconnected, planning a trajectory that minimizes the time to restore connectivity is challenging, particularly when the channel is not fully known.

statistics of the underlying channel. Further suppose there exists a known, non-empty set of points,  $Q_f$ , where, with probability 1, the channel is sufficiently strong to allow for communication when the UV is at any of those points. The problem of minimizing the distance traveled to connectivity can be posed as follows:

$$\min_{\{q[k]\}_{k=1, K}^K} \mathbb{E} \left[ \sum_{k=0}^{k_{\text{fpd}}-1} \|q[k] - q[k+1]\|_2 \right] \quad (22a)$$

$$\text{s.t.} \quad \text{Constraints (21d), (18),} \\ q[K] \in Q_f \quad (22b)$$

where  $k_{\text{fpd}}$  is defined as  $\min\{k \mid (21c) \text{ holds}\}$ , and, as  $k_{\text{fpd}}$  is a random variable due to imperfect channel information, the expectation is taken with respect to  $k_{\text{fpd}}$ . Importantly, Constraint (22b) ensures the expectation is always well-defined.

The problem of minimizing the expected distance to connectivity in a general sense is studied in [181]. A graph-theoretical formulation is put forth, where at each node, the UV has some probability of success, and the UV must plan a path through the graph which minimizes the expected distance traveled until success occurs. A Markov decision process describes the problem, which is proven to be NP-hard. To find high-quality approximate solutions, the problem is recast as a potential game with the nodes acting as the players, and log-linear learning is used to find near-optimal solutions. Results show that first seeking out nearby locations with even modest probabilities of connectivity can reduce average distance traveled compared to moving directly to the nearest connected point.

When the UV's trajectory is given beforehand, [182] provides rigorous analysis of the first-passage distance (FPD), the expected distance the UV will travel before reconnecting. An Ornstein-Uhlenbeck [200] process is shown to model log-normal channel shadowing along a straight line path, which permits an analytical expression for the FPD. If small-scale fading is not negligible, a recursive formulation is developed that allows the FPD to be calculated over a discretization of the path.

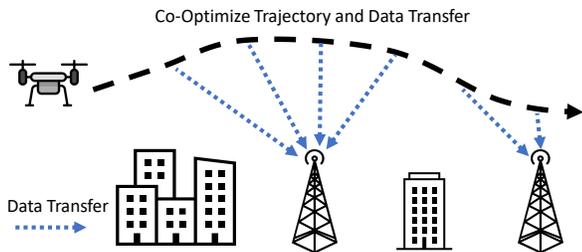


Fig. 11. Scheduling of wireless data offloading from a UV to a BS with complementary trajectory design can be a challenging problem, but it can also significantly reduce the energy needed for data transfer.

**Takeaways:** Quickly finding connected points is critical for both the safety and performance of the UV. These problems have a similar shortest-path element as those in (21), but when accounting for shadowing and path loss statistically, they become *stochastic* shortest path problems [110], which are much more difficult to solve.

3) **Scheduling Wireless Data Uploading:** In a number of scenarios, UVs must offload data to a remote station while traveling. In these situations, the UV should jointly optimize its trajectory and the corresponding adaptive transmission power/rate over the whole trajectory, for efficient data transfer. The general problem scenario is shown in Figure 11.

A core problem involves a UV moving from  $q_o$  to  $q_f$ , while offloading  $D_0$  bits of data along the way. The problem can be formulated starting with Problem (21) and replacing the SNR constraint (21c) with a data offloading constraint:

$$D[0] = D_0, \quad (23a)$$

$$D[k] = \max(D[k-1] - \Delta t B r(P[k], q[k]), 0), \quad (23b)$$

$$D[K] = 0, \quad (23c)$$

where  $B$  is the bandwidth of the communication channel,  $r(P, q)$  gives the data rate between the UV and the BS when the UV transmits with power  $P$  from position  $q$ . Alternatively, an additional term capturing the timeliness of data offloading can be added to the objective.

Scheduling when to upload data adds complexity, particularly for optimization of the transmit power. In problem (21), it is easy to see that transmit power should be chosen to satisfy Constraint (21c) with equality, so that the choices of transmit power,  $P[k]$ , are not directly coupled over time. However, the introduction of Constraint (23) results in tighter coupling of transmit power across time. We next present interesting studies related to this problem.

For a UV moving along a predefined trajectory, [183] considers jointly planning data transmission (*i.e.*, communication rate/power) and velocity in order to minimize total energy costs. It is shown that the UV should move quickly through areas of poor communication channel quality and more slowly through regions of high channel quality (via an optimized velocity profile), so that it spends more time transmitting in more favorable areas. Using the channel prediction framework of [59], an adaptive, online strategy is proposed. The proposed method significantly reduces energy consumption compared to

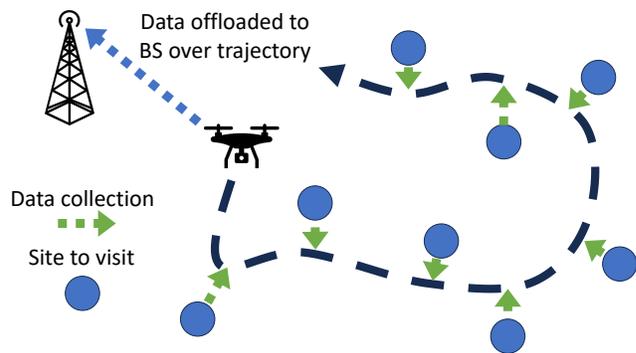


Fig. 12. Site visiting problems arise in a number of situations, including surveillance, surveying, and IoT/WSN systems. In addition to collecting data from the sites, the UV may also need to transfer the data back to a remote BS in a timely manner.

the baseline where the UV moves with constant speed and transmits with constant spectral efficiency.

A continuous time, energy-minimizing version of the problem appears in [184], where it is reformulated by moving Constraints (21e) and (23c) into the objective as additional penalty terms. This makes it in the form of an abstract Bolza optimal control problem [201]. This work is extended to a multi-UV case in [185], which proposes a distributed version of the optimal control solution. Results show that the UVs seek out areas of better channel quality and perform most of the wireless data uploading there.

**Takeaways:** Jointly adapting communication rate and UV speed to channel conditions can greatly improve communication energy efficiency. This may motivate planning trajectories that seek out areas of better channel quality, though the associated motion energy expenditure must also be weighed as well.

4) **Site-Visiting Problems:** As a final set of problems, we consider scenarios where the UV visits several sites to collect data either with onboard sensors, such as cameras, or through communication channels, *e.g.*, in the case of data collection from wireless sensor nodes or IoT devices. Additionally, the problem may involve transferring the collected data to a remote BS, either over a wireless communication link, in which case the UV's role resembles that of a relay, or by physically moving the data to an access point, sometimes referred to as data muling [202]. A general problem scenario is illustrated in Figure 12.

Usually, site visiting constraints ensure that all sites are visited, and the visit order must be determined, which introduces a combinatorial aspect to these problems. Site visiting constraints can take many forms. At the simplest, they may constrain the trajectory to pass through a (possibly singleton) set  $Q_\omega$ , *e.g.*, there exists a point in the trajectory,  $q[k]$ , where  $q[k] \in Q_\omega$  for all  $\omega$  in the set of sites  $\Omega$ . When sites correspond to the locations of IoT devices or sensor nodes, the visiting constraint may ensure that a set amount of data is uploaded to the UV.

Site visiting problems also bring with them a distinct set of metrics, which may enter the system either as constraints or objectives. When the sites correspond to communication

nodes, the energy required to upload the data to the UV is an important measure of system performance, as these nodes are frequently quite energy-limited. Additionally, metrics such as average delay and Age of Information (AoI) [203] capture the timeliness of data transfer to the BS.

A general site visiting problem can be expressed as:

$$\begin{aligned} & \min_{\substack{\{q[k]\}, \mathcal{R}_{UV}, \\ \mathcal{R}_\Omega}} f_{\text{GLB}}(\mathcal{E}_{UV}, \{\mathcal{E}_\omega\}_{\omega \in \Omega}, f_T(\{q[k]\}, \mathcal{R})) \\ & \text{s.t. (18) Kinematic constraints,} \\ & \quad g_c(\{q[k]\}, \mathcal{R}) \leq 0 \text{ Comm. res. constraints,} \\ & \quad \left. \begin{aligned} g_D(\mathcal{R}, \{q[k]\}) \leq 0 \\ h_D(\mathcal{R}, \{q[k]\}) = 0 \end{aligned} \right\} \begin{array}{l} \text{Site Visiting} \\ \text{Constraints} \end{array}, \end{aligned} \quad (24)$$

where  $\mathcal{E}_{UV}$  is the total energy consumed by the UV,  $\mathcal{E}_\omega$  is the energy expended by site  $\omega$  to upload data to the UV,  $\mathcal{R}_{UV}$  and  $\mathcal{R}_\Omega$  are the communication resource variables for the UV and sites, respectively, with  $\mathcal{R} = \mathcal{R}_{UV} \cup \mathcal{R}_\Omega$ , and  $f_T(\{q[k]\}, \mathcal{R}_{UV}, \mathcal{R}_\Omega)$  is a measure of the timeliness of data collection, such as mission completion time or AoI.

**Data Collection for IoT/WSN:** A number of works focus on the problem of data collection from IoT/WSN devices without explicitly modeling the data transfer from UV to BS. In these cases, the energy consumed by the devices,  $\mathcal{E}_\omega$  is often given greater attention. For example, [193] studies the problem of deploying multiple UAVs to collect data from IoT devices in a way that minimizes energy expenditure by the devices. Using BCD and sequential quadratic programming, the authors find device-to-UAV associations, UAV position, and device transmit power schedules that greatly reduces the total energy consumed by the IoT devices compared to the scenario where the UAVs are statically placed in the environment.

In [194], the authors consider the challenging problem of multiple UAVs collecting data from IoT devices under time constraints, that is, when the devices are only awake for a subset of the time. They formulate a mixed-integer nonlinear program (MINLP) to find a trajectory and an allocation of communication resource which maximize the total number of IoT nodes serviced over a finite time frame. A similar time-constrained, multi-UAV problem is considered in [195], which instead focuses on the problem of minimizing the total UAV energy consumption while ensuring that a set amount of data is collected from each node. The problem is solved using results from the literature on capacitated vehicle routing problems with time windows (CVRPTW).

**Sensing and Data Transfer:** A different set of problems focus on the UV-to-BS offloading aspect of these problems. These problems are similar to the data offloading problem in Section IV-B3, but with modified data dynamics. Specifically, after visiting each site, an additional amount of data enters the UVs onboard memory, so that Eq. (23b) is replaced by

$$D[k] = \max\left(D[k-1] + \sum_{\omega \in \Omega} I_\omega[k] D_\omega - \Delta t B r(P[k], q[k]), 0\right),$$

where  $I_\omega[k]$  indicates if site  $\omega$  is visited at time step  $k$  and  $D_\omega$  is the data associated with site  $\omega$ .

In [191], a UV visits a number of sites where it collects data and transmits them to a remote station. Initially, the problem is considered in the absence of the sites, and the UV simply decides whether it is more energy efficient to transmit from the current location or to first expend energy to move to another location and then transmit. Then, the problem of minimizing energy consumption by co-optimizing the UV's motion and communication strategies for retrieving and relaying data from the site back to a BS is formulated as a MILP.

A related problem is studied in [196], in which multiple UVs service a number of sites to keep a generic Quantity of Interest (QoI) from growing unbounded while maintaining communication with a remote BS. Two cases of connectivity are considered. In the first, the UVs must establish a communication link to a remote BS after each site visit. In the second, the UVs must only connect once during their cyclic trajectories. The problem of minimizing the total energy consumption while guaranteeing boundedness of the quantity as well as respecting the connectivity constraints is formulated as a MILP.

**Data Collection and Relay:** A final set of problems considers uplink communication both from the sites to the UV and from the UV to the BS. These types of problems are the most general and frequently feature metrics related to the timeliness of end-to-end data transfer. For example, [187] studies the problem of minimizing the average wait time in a data collection and relay scenario. The authors consider a single UV relaying data between a set of spatially diverse and disparate source-destination pairs, where heterogeneous data accumulation rates at the sources necessitate different visiting frequencies. They then find the optimal robotic routing policies (e.g., optimal relay locations as well as the sequence with which the pairs are serviced). More specifically, the problem of minimizing the average time a bit of data waits at the source before being relayed to the destination is approached from a polling systems perspective [204], and several fundamental properties of the system are proved using results from queuing theory, all the while showing how stochastic robotic routing can be handled.

Another metric of the timeliness of data transfer is Age of Information (AoI) [203], which measures the recency of the most up-to-date data at the BS. In [197], the authors look at the problem of minimizing the AoI in a scenario where multiple UAVs are dispatched to collect data from clusters of devices and then return to a data center to offload the data. The AoI accounts for the travel time as well as the device-to-UV and UV-to-BS communication times. The problem is approached by first partitioning the sites among the UAVs with a variant of K-means [205], then using ant colony optimization [206] to find the visit order for each UAV. For a single UAV, end-to-end AoI is also addressed in [198], but the data transfer to the BS is constrained to happen directly after receiving the data from the device, that is, the UAV operates as a half-duplex relay. The authors use BCD along with SCA and convex optimization methods to produce trajectories and transmit power profiles which minimize the average AoI while respecting communication energy constraints for the UAV and devices.

**Takeaways:** The combinatorial element of site visiting and wireless offloading with causality constraints together produce very complex problems with elements of both continuous and discrete optimization coupled strongly across time. Furthermore, when the UVs collect data via wireless uploading from devices, additional considerations come into play, including the energy consumed by the devices and possibly timing windows. For end-to-end data transfer with data generating dynamically at each site, the system may be viewed as a network of queues [207], where data accumulating in queues at the sensor nodes are uploaded to the queue on the UV which in turn uploads the data to a BS. The system also bears resemblance to a polling system, in which a single server serves multiple queues, with some idle time incurred when switching between the queues.

### C. Stochastic Dynamic Programming Formulations

Many of the analytical formulations of the previous section (along with the accompanying solution methods) account both for the inherent stochasticity of UV-assisted wireless systems and for the optimization of system parameters over an extended period of system operation. Nonetheless, the recent paradigm of DRL introduces a compelling framework for advancing these formulations. In a dedicated subsection, we delve into the evolving landscape of stochastic dynamic programming formulations within UV-assisted networking. Furthermore, we explore the application of learning-based methods (and more prominently DRL) to effectively address such settings.

1) *Unmanned Vehicles as Providers:* The current subsection provides an overview of UV problem formulations that adhere to the general framework of stochastic dynamic programming and consider that the UV acts as a provider of system services. In such role, the UV can be a central node that provides connectivity with the core network, for example it could be an aerial base station (BS). Alternatively, the UV can be acting as an intermediate/relay node that extends coverage and connectivity.

**UVs as Base Stations:** Assume a UV acts as an aerial BS and serves  $\mathcal{U}$  ground users, that are located at specific positions. The elevation level of the UV stays constant during the whole time horizon of operation. The UV starts at position  $q_0$  and is supposed to finish its flying trajectory at position  $q_K$  at time  $K$ . The channel at time  $k$  between the UV and the ground user  $i$ ,  $\mathcal{H}_i$ , depends on the position of the UV  $q[k]$ ,  $\mathcal{H}_i(q[k])$ . The transmission power of the UV at time step  $k$  is denoted as  $P[k]$ . The communication rate for user  $i$  at time  $k$  is  $\log_2(1 + \text{SNR}(\mathcal{H}_i(q[k])))$ . The set of all valid trajectories in the UV operation space that amount to overall time  $T$  is denoted as  $\mathcal{Q}$ .

$$\max_{\mathcal{Q}} \mathbb{E}_{\mathcal{Q}} \left[ \sum_{k=0}^K \sum_{i \in \mathcal{U}} \log_2(1 + \text{SNR}(\mathcal{H}_i(q[k]))) \right] \quad (25a)$$

$$\text{s.t. } q[0] = q_0, \quad (25b)$$

$$q[K] = q_K, \quad (25c)$$

$$q[k] \in [q_{min}, q_{max}]^2 \quad \forall k \in [0, K], \quad (25d)$$

$$\log_2(1 + \text{SNR}(\mathcal{H}_i(q[k]))) \geq \zeta_{min} \quad , \quad (25e)$$

$$\sum_{k=0}^K P[k] \leq P_{total} \quad (25f)$$

The goal is to design the UV trajectory over time such that the sum of total communication rate is maximized in expectation. The constraints correspond to the fact that the initial and final positions of the UV are predetermined, the UV area of motion is confined within a  $2D$  rectangular space, and there should be fairness among end users in the sense that the communication rate of any individual user is not below a certain threshold at all time steps. Finally, the overall power budget for transmission is below a certain predefined threshold.

The problem formulation calls for approaches where optimization unfolds through time. DRL is an organically suitable paradigm for tackling the problem. In order to employ DRL for settings that fall under the above formulation, the Markov decision process (MDP, see Section III-B) must be defined:

- **State:**  $s[k] = f_s(q[k])$ , the state is a vector that is a function of the UV position at time step  $k$ . In the simplest case, the state can be the position itself, or a window of previous positions.
- **Action:**  $a[t] = f_a(\delta q[k])$ . The action is a function of the displacement of the UV for time step  $k$ . The displacement can be a vector that represents the difference between the position at the current step and the position in the subsequent step. The range of the displacement should adhere to the imposed kinematic constraints.
- **Reward:**  $r[k] = \sum_{i \in \mathcal{U}} \log_2(1 + \text{SNR}(\mathcal{H}_i(q[k+1])))$ . The reward is the overall communication rate achieved at time step  $k+1$ . One should note that, by the definition of the MDP, the reward is calculated after performing the action at the current state. In order to facilitate the fairness constraint, the reward can be augmented by the difference between the minimum communication rate among end users at each time step and the minimum acceptable communication rate  $\zeta_{min}$  and the transmission power at every time step.

A significant body of work exists that falls under the aforementioned formulation. The work in [139] examines a scenario where a UV acts as a BS and the goal is to optimize the total communication rate over time. A Q-learning [208] approach is considered and the space of UV motion is discretized into a fine grid. In [140], the authors focus on a similar scenario with multiple UV-BSs, where the goal is to maximize a weighted throughput objective that relates to fairness among end users. A Multi-agent Deep Deterministic Policy Gradient approach is proposed that disentangles the contributions of the individual

UVs. The authors in [144] examine a setting where a UV acts as a BS serving end users in the context of Open Radio Access Networks [209] under MIMO fading channels. The UV motion is discretized, and two RL approaches are proposed to maximize the network's throughput over time. The first approach is a variant of SARSA [210] and the second approach is predicated on offline Q-learning [211].

The work of [141] examines the trajectory optimization of multiple UVs that serve a ground cellular network. The goal is to design the trajectories of the UV swarm and the task completion time in order to minimize energy expenditure, subject to a constraint that pertains to the continuous connectivity of the UVs with the ground cellular network. The authors consider two scenarios: one where the radio channel maps are known to the swarm and one where the channel maps are initially unknown. For the scenario where the channel maps are known, they employ a convex optimization approach along with a dynamic-weight shortest path algorithm. For the second scenario, the aforementioned approach is non-applicable, therefore the authors propose a deep Q-learning approach with a dueling architecture in order to optimize the UV trajectories over time. The work of [142] examines a multi-UV cellular network, where the UV trajectories are optimized with a TD-learning [212] approach in order to minimize the overall task completion time. The state-action value function is linear and tile-coding [213] is implemented in order to ameliorate the effect of the large state-action space. A similar scenario is considered in [143], with the difference being that only one UV is considered. A Deterministic Policy Gradient algorithm is developed to optimize the UV motion with the goal of maximizing the uplink sum-rate.

**Cooperative Communication:** Besides the UV operating as a BS, there has been increasing interest in applications where the UVs are deployed as intermediate nodes in the cascade of the network [214]. The inherent mobility of UVs can be leveraged in order to increase coverage and improve the QoS on the receiver side.

Let us consider  $I$  UVs acting as relays that facilitate the communication between a source  $S$  and a destination  $D$ . The locations of source and destination are fixed throughout the operation, therefore the dependence on their location is dropped from all expressions. The source transmits symbol  $s[k] \in \mathbb{C}$  with power  $P$  and the signal received at UAV/relay  $i$  which is at position  $q_i[k]$ ,  $x_i[q_i[k], k]$  depends on the position of the relay and the channel from the source to the relay/UV. The channel from the source to the relay,  $\mathcal{H}_{S_i}$  also depends on the relay position and possible on the time (assuming spatio-temporally correlated channels [82]). The signal received by relay  $i$ ,  $x_i[k]$  can be generally expressed as:

$$x_i[k] = x_i(q_i[k], \mathcal{H}_{S_i}[k]). \quad (26)$$

When considering narrowband flat fading channels, the effect of the channel on the transmitted signal is multiplicative [82], [127]. On the other hand, wideband channels of 6G systems are frequency selective [215], [216] (refer to Section II-B).

Each of the UVs/relays multiplies its received signal with weight  $w_i[k]$  and all UVs collectively transmit their signals to the destination using a collective power budget  $P_T$ . The

channel from each UV to the destination,  $\mathcal{H}_{iD}$ , depends on the UV's position and the time instance. The received signal at the destination  $x_D$  is a function of all channels (from the source to the UV and from the UV to the destination) and all relay positions:

$$x_D[k] = \mathcal{F}(\{q_i[k]\}_{i=1}^I, \{\mathcal{H}_{S_i}\}_{i=1}^I, \{\mathcal{H}_{iD}\}_{i=1}^I, k), \quad (27)$$

where  $\{\mathcal{H}_{S_i}\}_{i=1}^I$  denotes the channels from the source to all deployed UVs/relays and  $\{\mathcal{H}_{iD}\}_{i=1}^I$  the channels from all relays to the destination. Noise is induced both at the signal reception at each UV and at the signal reception at the destination. Therefore, the destination signal can be decomposed as:

$$x_D[k] = x_D^{signal}[k] + x_D^{noise}[k], \quad (28)$$

Under the assumption that the channels (both from the source to each UV and from each UV to the destination) exhibit correlations with respect to time and space, the problem can be formulated as:

$$\max_{\mathcal{Q}} \mathbb{E}_{\mathcal{Q}} \left[ \sum_{k=0}^{\infty} \mathcal{U}(\{\mathcal{H}_{S_i}[k]\}_{i=1}^I, \{\mathcal{H}_{iD}[k]\}_{i=1}^I) \right] \quad (29a)$$

$$\text{s.t. } q_i[0] = q_{i0} \quad \forall i, \quad (29b)$$

$$\|q_i[k+1] - q_i[k]\|_2 \leq v_{\max} \Delta_k \quad \forall i, k, \quad (29c)$$

$$\mathbb{E}[\|x_i[q_i[k], k]\|^2] \leq P_I \quad (29d)$$

The set of all infinite duration permissible trajectories of UVs is denoted as  $\mathcal{Q}$ .  $\mathcal{U}$  is a general utility function of the system that depends on the realization of the channels and therefore on the corresponding relay positions for all time steps of operation. The most typical choice for the utility is the SINR at the destination, which is expressed as:

$$\mathcal{U}[k] \equiv \text{SINR}[k] = \frac{|x_D^{signal}[k]|^2}{|x_D^{noise}[k]|^2}, \quad (30)$$

Besides the SINR, there exists a plethora of other candidate utility metrics, such as the sum-rate (for multiple receivers) and the secrecy rate (refer to Section II-C2). The choice of the utility function introduces different nuances that are domain-specific and have to be explicitly addressed (e.g., additional interference when considering multiple destinations). The above formulation paves the way for DRL approaches to be adopted. The works [152], [217], [218] consider multiple UVs as relays facilitating the communication of a single source-destination pair. The space over which the UVs can move is discretized into 2D grid cells, and various variations of the deep Q-learning approach [117], [129] are proposed to solve the problem. In those works, the authors noted the inability of Multilayer Perceptrons (MLPs) with ReLU activations to capture the high frequency spectral components of the optimal value function, which arise due to the pronounced variability of the underlying channels with respect to time and space. It was also noted that by preprocessing the input of the state of the corresponding MDP with a learnable Fourier kernel, those issues can be mitigated, leading to an increase in convergence speed and average SNR accumulation at the receiver. In [153], the assumption for space discretization is

relaxed and a variation of the Soft Actor-Critic [124] algorithm for continuous motion control is proposed.

The works [75], [147] also examine the aforementioned problem and propose a stochastic programming approach to solve it. The approach is data-driven and optimizes the SNR in expectation for the subsequent time step of system operation. In [219], the authors perform a comparative analysis of the performance of the DRL approach of [152] and the stochastic programming approach of [75]. Finally, the work [128] considers the same problem and develops a multi-armed bandit method to tackle it.

The work [159] studies a scenario where a single UV serves multiple users as a relay. Multiple model-free RL algorithms are proposed for UV motion control with the goal of maximizing energy efficiency under quality of service constraints. The proposed RL approaches are shown to surpass the performance of a connected hybrid precoding approach. The work of [220] examines scenarios where a UAV serves as a mobile relay in order to maintain the communication between multiple destinations (end users) and a BS. Both model-free and model-based RL approaches are proposed for the same setup. In [155], the authors examine a scenario where one UAV acts as a relay and propose an on-policy DRL approach for trajectory design with the goal being to optimize the average throughput of the collective system. The DRL approach is a proximal policy optimization method [121]. The proposed algorithm accounts for the molecular absorption effect that is prominent in THz communications [221].

The work of [222] examines a communication set up where multiple UVs serve as relays in order to facilitate the communication between a BS and multiple end users. A deep Q-learning approach with a dueling architecture of the Q-network [129] is proposed in order to optimize the UV locations over time. The goal is to design the trajectories of all deployed UVs in order to maximize the downlink capacity of the network under the constraint that all end users should be served by at least one relay at all times of system operation.

Both this subsection and the previous one offer insights and problem formulations concerning the intersection between DRL and UV motion control. The distinction between the two lies in the specific roles played by UVs in each scenario. In the first subsection, the UV operates as a provider, akin to a mobile BS, while in the second subsection, it acts as an intermediate mobile node facilitating communication between source and destination points.

At a cursory glance, the formulations presented in Problem (25) and the Problem of (29) might appear similar. Both require the design of UV motion policies to maximize the expected sum of a QoS metric, contingent on the channel conditions, over an extended time horizon. Solving sequential decision-making problems using DRL hinges significantly on estimating the optimal value function of the corresponding MDP. In both cases, this value function depends on the underlying channels. When the UV serves as a provider, only the channel from the UV to the user impacts the value function. Conversely, when the UV acts as a relay or helper, the optimal value function is influenced by both the channels from the UV to the receiver and the channels from the source or BS to the

UV. This additional dependency introduces localized variability in the true value function, making its accurate estimation through bootstrapping a challenging endeavor [127].

2) *Unmanned Vehicles as Consumers*: This subsection delves into the collaborative relationship between DRL and UVs specifically when UVs serve as consumers of services within wireless networks. DRL is commonly used for UV motion control or resource allocation, especially in “data harvesting” scenarios from IoT/sensor networks. The inherent mobility of UVs is advantageous for efficient data collection from dispersed sensors. However, challenges arise, requiring advanced algorithms to optimize UV motion for objectives like maximizing data throughput, operational time, data freshness, and minimizing energy consumption amidst environmental randomness.

A broad scenario that encompasses various settings of sequential decision-making for UV-assisted data collection unfolds as follows: A single UV is tasked with collecting data from  $M$  clusters of sensors, each cluster conveniently comprising the same number of  $I$  sensors. The UV’s position at each step is represented as  $q[k] = [x[k], y[k], z[k]] \in \mathbb{R}^3$ . The position of the  $i$ th sensor within the  $m$ th cluster at time step  $t$  is denoted as  $q_i^m = [x_i^m[k], y_i^m[k]]$ . The throughput from sensor  $i$  within cluster  $m$  depends on both the UV’s position and the sensor’s location, under particular assumptions for the underlying channel model:  $R_{i,m}[k] = R_{i,m}[q[k], q_i^m[k]]$ . Considering the existence of a probability of link blockage between the UV and each sensor in every cluster, the sensor’s throughput can be defined as  $R_{i,m}[k] = P_{i,m}[k]R_{i,m}[q[k], q_i^m[k]]$ , where  $P_{i,m}[k] \in [0, 1]$  signifies the probability of a lost connection between the UV and the  $i$ th user within the  $m$ th cluster at time step  $k$ . The overarching aim of the sequential decision-making framework is to construct the policy for the UV motion such that the overall throughput achieved throughout the mission is maximized while simultaneously the overall mission time is minimized.

On denoting the set of all permissible UV trajectories by  $\mathcal{Q}$ , the generic optimization problem can be defined as

$$\max_{\mathcal{Q}} \mathbb{E}_{\mathcal{Q}} \left[ \sum_{k=0}^K \sum_i^I \sum_m^M \alpha R_{i,m}[q[k], q_i^m[k]] - \beta k \right] \quad (31a)$$

where  $\alpha$  and  $\beta$  are scalars that manage the design tradeoff between the desire for overall sum-rate maximization and mission completion time minimization.

The work presented in [186] considers a scenario very similar to the one described above. The state is the position of the UV at each time step, the operational space is discretized into a fine grid and the action space is comprised by the displacement of the UV to one of the neighboring cells. The authors propose the adoption of the deep Q-learning algorithm to control the UV positions over time in order to optimize a linear combination of the overall sum-rate and the mission completion time. The study presented in [189] considers a scenario where multiple UVs perform both data collection and wireless energy transfer. In particular, each UV supplies energy to each one of the sensors. Subsequently, the sensors transmit their collected data to the UVs. However, when the same UV is used both for data collection and

energy harvesting, both operations are negatively affected. The authors propose splitting the group of UVs into two categories, one that is solely responsible for energy harvesting and another that is solely responsible for data collection. The goal is to optimize the trajectories of both groups of UVs. Since the optimal policies of the two groups are correlated, the authors propose a multi-agent DRL algorithm in order to control the trajectories of the UVs of each category such that a linear combination of the overall throughput, the Age-of-Information, the overall harvested energy and the UV energy expenditure is optimized. The authors of [190] examine a challenge, particular to UV-assisted data collection in IoT systems, that pertains to the unsuccessful data transmission due to buffer overflow caused by the maneuvering of UVs. Since the state of the individual IoT nodes can be potentially obscure to the UV, the induced MDP is Partially Observable [223]. The authors propose a deep Q-learning-based approach for Partially Observable MDPs in order to control both the UV flight and the data scheduling with the goal of minimizing the overall packet loss.

A similar data collection problem is considered in [192], which additionally considers the data transfers from the UV to a remote station. The UV's trajectory is co-optimized with an adaptive transmit rate for wireless data transfer to a remote station, and an MDP is formulated, with a cost function which captures both the motion and communication energy of the UV, and a provably-optimal Monte-Carlo tree search (MCTS) method is proposed which treats the optimization of communication resources along the entire path as the final stage in the MDP, thus greatly reducing the size of the state space. In a related work, [188] examines the persistent dynamic relay scenario first studied in [187], in which a single UV services a group of disparate source-destination pairs. The authors are then interested in finding the optimal robotic routing locations and dynamic routing policies. The authors show that the problem of minimizing the average wait time can be posed as an average-cost semi-Markov decision process, and propose using a variation of PPO to efficiently solve the problem, while showing how to handle a long-run average cost criterion.

## V. ADDITIONAL CONSIDERATIONS

### A. Security

The affordability of hardware and the high probability of LoS communication links make UVs appealing for wireless communications, yet they can also present vulnerabilities. This becomes particularly crucial in military applications, where UVs often carry sensitive information, making privacy and security paramount requirements of system design.

One challenge lies in the low-cost hardware mounted on UVs, which may struggle to run sophisticated encryption algorithms necessary for ensuring information security. As a result, UVs become susceptible to cybersecurity attacks. These threats may seek to disrupt access to critical services, as in Denial-of-Service (DoS) [224] and battery attacks [225], or they may compromise the privacy of sensitive information, as in spoofing attacks [226].

However, even if computational resources are available, traditional cryptographic approaches may be difficult to implement on a UV. This is because they rely on secret keys, which are difficult to distribute in a wireless network. Physical layer security (PLS) approaches may thus be more applicable in such scenarios. By exploiting the physical characteristics of the wireless channel, PLS design aims to enable the legitimate destination to obtain the source information successfully, while preventing an eavesdropper (ED), or unauthorized user from decoding the information [227]. PLS system design aims to maximize the secrecy rate, which is the rate the legitimate receiver receives minus the rate that the eavesdropper receives, subject to certain system constraints. PLS design of communication systems has been well investigated [228]–[234]. One approach to ensure PLS is *cooperative jamming*, where trusted relays act as helpers and beamform artificial noise (AN), aiming to degrade the ED's channel [228]–[232]. Another approach is for the source to beamform AN along with the intended information for users, in a way that the users do not experience interference [233], [234]. PLS design for DFRC systems has been considered in [235], [236], while the design of DFRC aided by an RIS is investigated in [237].

### B. Edge Computing

UVs can benefit from [238] and contribute to edge computing and edge intelligence [239]. The effective functioning of UVs, both as consumers and providers of wireless connectivity, relies on the application of computationally intensive algorithms, such as vision algorithms for precise localization and mapping [240]. The traditional approach of offloading such operations to a central cloud server is impractical due to latency issues that exceed stringent requirements [238]. Therefore, leveraging edge servers near UVs allows for substantial storage and computation, striking a balance between latency and computational capabilities. Particularly in the realm of artificial intelligence on the edge, various research areas such as transfer learning [241], knowledge distillation [242], and neural network pruning/quantization [243] are emerging to facilitate processing on the network edge, in close network proximity to the UVs.

Conversely, UVs can serve as edge nodes where end users offload tasks [244]. The mobility of UVs introduces a dynamic edge computing environment, allowing them to approach users based on a flexible schedule. This dynamic interaction reduces latency and enhances throughput for end users in response to evolving network conditions. Additionally, edge computing brings indirect security benefits to UV operation. As UVs are equipped with onboard sensors such as cameras, they inherently collect data that is private to end users. Leveraging the paradigm of federated learning [245] alongside edge computing helps avoid the transfer and storage of sensitive information on central cloud servers. Instead, storage and computation are offloaded to edge nodes, which are less susceptible to information interception and, therefore, invasion of privacy.

### C. Federated Learning

Undoubtedly, the integration of UVs into wireless networking is closely associated with the rise of deep learning methods [246]. These approaches prove valuable for optimizing components of the UV system configuration and harnessing the vast amount of data collected or generated by UVs. Deep learning paradigms, known for their data-intensive nature, demand substantial computational resources, typically provided by cloud servers. However, this gives rise to two significant concerns: Firstly, transferring data from UVs to the central cloud introduces latency, which can be prohibitive for many applications. Secondly, the transmitted data may include sensitive information, posing privacy and security risks.

The federated learning paradigm [247] addresses these concerns by keeping data at the terminal device that generated or collected it [248]. In this approach, a local deep learning model is trained by the data, and only the parameters of the local model are transmitted to the central cloud center. This methodology helps mitigate privacy invasion for end-users. In addition, the transmission of model parameters induces less latency than the transmission of raw data. A plethora of different works leverage federated learning for UV-assisted applications. In [249], the authors examine distributed image classification for multi-UV exploration. The authors in [250] propose a federated learning scheme that addresses nuanced privacy requirements in UV-assisted crowdsensing. Overall, utilizing federated learning for UV-assisted 6G networks remains an open and exciting future research direction.

## VI. CONCLUSIONS

The tech world eagerly anticipates the advent of the 6th Generation of wireless networking, promising transformative capabilities and applications. The success of this revolution, from closing of the digital divide to fueling the industrial sector, hinges on diverse technological advancements. Unmanned Vehicles emerge as a pivotal technology, especially crucial for extended coverage and connectivity. While UVs enhance wireless infrastructure, they also introduce novel challenges requiring attention from academia and industry.

This paper offers a comprehensive view of settings and scenarios at the intersection of UVs and wireless networking, accompanied by an in-depth exploration of corresponding problem formulations and methods. To present this information effectively, the central thrust of the paper has been the unification of the problem space, providing a cohesive framework to understand the use cases, problem formulations, and necessary tools. Overall, the paper can equip researchers, engineers, and stakeholders with a clear understanding of both the opportunities and challenges inherent in the convergence of unmanned vehicles and 6G networks, paving the way towards a successful integration.

## REFERENCES

[1] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021.

[2] 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, 2020.

[3] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, 2020.

[4] "Framework and overall objectives of the future development of IMT for 2030 and beyond," ITU-R, Tech. Rep., 2023, draft New Recommendation.

[5] A. Chaoub, M. Giordani, B. Lall, V. Bhatia, A. Kliks, L. Mendes, K. Rabie, H. Saarnisaari, A. Singhal, N. Zhang, S. Dixit, and M. Zorzi, "6G for bridging the digital divide: Wireless connectivity to remote areas," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 160–168, 2022.

[6] "6G: The Next Hyper-Connected Experience for All," Samsung, Tech. Rep., 2020.

[7] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, 2019.

[8] "The ITU-R Framework for IMT-2030," ITU-R, Tech. Rep., 2023. [Online]. Available: [https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2030/Documents/IMT-2030%20Framework\\_WP%205D%20Management%20Team.pdf](https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2030/Documents/IMT-2030%20Framework_WP%205D%20Management%20Team.pdf)

[9] "Recommendation ITU-R M.2083-0: IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond," ITU-R, Tech. Rep., 2015.

[10] R. Liu, R. Y.-N. Li, M. D. Renzo, and L. Hanzo, "A vision and an evolutionary framework for 6G: Scenarios, capabilities and enablers," 2023.

[11] C. de Alwis, Q.-V. Pham, and M. Liyanage, "Smart Cities and Society 5.0," in *6G Frontiers: Towards Future Wireless Systems*, 2023, pp. 197–210.

[12] S. Polymeni, S. Plastras, D. N. Skoutas, G. Kormentzas, and C. Skianis, "The impact of 6G-IoT technologies on the development of agriculture 5.0: A review," *Electronics*, vol. 12, no. 12, 2023. [Online]. Available: <https://www.mdpi.com/2079-9292/12/12/2651>

[13] H. Shin, S. Park, L. Kim, J. Kim, T. Kim, Y. Song, and S. Lee, "The future service scenarios of 6G telecommunications technology," *Telecommunications Policy*, p. 102678, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0308596123001891>

[14] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[15] P. R. Singh, V. K. Singh, R. Yadav, and S. N. Chaurasia, "6G networks for artificial intelligence-enabled smart cities applications: A scoping review," *Telematics and Informatics Reports*, vol. 9, p. 100044, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S277250302300004X>

[16] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, 2019.

[17] H. Sun, C. Ng, Y. Huo, R. Q. Hu, N. Wang, C.-M. Chen, K. Vasudevan, J. Yang, W. Montlouis, D. Ayanda, K. V. Mishra, K. Tekbryik, N. Hussain, H. K. Sahoo, and Y. Miao, "Massive MIMO," in *2022 IEEE Future Networks World Forum (FNWF)*, 2022, pp. 1–51.

[18] A. Pallaprolu, W. Hurst, S. Paul, and Y. Mostofi, "I beg to diffract: RF field programming with edges," in *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3570361.3613266>

[19] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, 2022.

[20] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, 2019.

[21] M. Di Renzo and A. I. Aravanis, "Catching the 6G wave by using metamaterials," in *Shaping Future 6G Networks: Needs, Impacts, and Technologies*, 2022, pp. 69–87.

[22] U. Ghafoor, M. Ali, H. Z. Khan, A. M. Siddiqui, and M. Naeem, "Noma and future 5G & B5G wireless networks: A paradigm," *Journal of Network and Computer Applications*, vol. 204, p. 103413, 2022.

- [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1084804522000728>
- [23] B. Clerckx, Y. Mao, E. A. Jorswieck, J. Yuan, D. J. Love, E. Erkip, and D. Niyato, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1265–1308, 2023.
  - [24] S. Stańczak, A. Keller, R. L. Cavalcante, N. Binder, and S. Velayutham, "Long-term Perspectives," in *Shaping Future 6G Networks: Needs, Impacts, and Technologies*, 2022, pp. 235–254.
  - [25] Z. M. Miličević and Z. B. Bojković, "From the early days of unmanned aerial vehicles (UAVs) to their integration into wireless networks," *Vojnotehnički glasnik/Military Technical Courier*, vol. 69, no. 4, pp. 941–962, 2021.
  - [26] A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5g network with the advent of 6G: Architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67 512–67 547, 2020.
  - [27] R. Shrestha, R. Bajracharya, and S. Kim, "6G enabled unmanned aerial vehicle traffic management: A perspective," *IEEE Access*, vol. 9, pp. 91 119–91 136, 2021.
  - [28] R. Bajracharya, R. Shrestha, S. Kim, and H. Jung, "6G nr-u based wireless infrastructure UAV: Standardization, opportunities, challenges and future scopes," *IEEE Access*, vol. 10, pp. 30 536–30 555, 2022.
  - [29] M. Mozaffari, X. Lin, and S. Hayes, "Toward 6G with connected sky: UAVs and beyond," *IEEE Communications Magazine*, vol. 59, no. 12, pp. 74–80, 2021.
  - [30] E. B. Parker, "Closing the digital divide in rural america," *Telecommunications Policy*, vol. 24, no. 4, pp. 281–290, 2000.
  - [31] G. Karabulut Kurt, M. G. Khoshkholgh, S. Alfattani, A. Ibrahim, T. S. J. Darwish, M. S. Alam, H. Yanikomeroglu, and A. Yongacoglu, "A vision and framework for the high altitude platform station (HAPS) networks of the future," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 729–779, 2021.
  - [32] A. Weissberger, "AT&T introduces 5G flying COWs (Cell on Wings) drones," November 2023, [Online; posted 22-June-2022]. [Online]. Available: <https://techblog.comsoc.org/2022/06/22/att-introduces-5g-flying-cows-cell-on-wings-drones/>
  - [33] S. Halawi, "Zephyr's next chapter," Airbus HAPS Connectivity Solutions, Tech. Rep., 2022, [https://mediaassets.airbus.com/pm\\_38\\_602\\_602492-1e1gtafyip.pdf](https://mediaassets.airbus.com/pm_38_602_602492-1e1gtafyip.pdf) [Online].
  - [34] "3GPP Release 17 - Building Blocks for UAV Applications," 3GPP, Tech. Rep., 2022.
  - [35] J. Stanczak, "NR Support for UAVs," August 2023, [Online; posted 06-July-2023]. [Online]. Available: <https://www.3gpp.org/technologies/nr-uav>
  - [36] "FAA unmanned aircraft systems integration pilot program," U.S. DoT, FAA, Tech. Rep., 2020.
  - [37] "BEYOND," August 2023, [Online]. [Online]. Available: [https://www.faa.gov/uas/programs\\_partnerships/beyond](https://www.faa.gov/uas/programs_partnerships/beyond)
  - [38] "Wing," <https://wing.com>, accessed: 2023-11-30.
  - [39] C. Chen, "11 photos of amazon's new prime air drone that can fly in light rain and deliver packages up to 5 pounds in under an hour," 2023, <https://www.aboutamazon.com/news/transportation/amazon-prime-air-drone-delivery-mk30-photos>.
  - [40] "Nuro," <https://nuro.ai>, accessed: 2023-11-30.
  - [41] S. Dresser, "Amazon announces 2 new ways it's using robots to assist employees and deliver for customers," 2022, <https://www.aboutamazon.com/news/operations/amazon-introduces-new-robotics-solutions> [Online].
  - [42] H. Shakhathreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48 572–48 634, 2019.
  - [43] "Autopilot and Full Self-Driving Capability," <https://www.tesla.com/support/autopilot>, accessed: 2023-11-30.
  - [44] "Waymo," <https://waymo.com>, accessed: 2023-11-30.
  - [45] "Autonomous mobility and delivery at Uber," <https://www.uber.com/us/en/autonomous>, accessed: 2023-11-30.
  - [46] Y. Mostofi, "Communication-aware motion planning in fading environments," in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 3169–3174.
  - [47] Y. Mostofi, M. Malmirchegini, and A. Ghaffarkhah, "Estimation of communication signal strength in robotic networks," in *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Anchorage, AK, USA, May 2010, pp. 1946–1951.
  - [48] A. Muralidharan and Y. Mostofi, "Communication-aware robotics: Exploiting motion for communication," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 115–139, 2021.
  - [49] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241–2263, 2019.
  - [50] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.
  - [51] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.
  - [52] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6G wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 122–134, 2020.
  - [53] Q. Wu, J. Xu, Y. Zeng, D. W. K. Ng, N. Al-Dahir, R. Schober, and A. L. Swindlehurst, "A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence," *IEEE J. Sel. A. Commun.*, vol. 39, no. 10, p. 2912–2945, oct 2021. [Online]. Available: <https://doi.org/10.1109/JSAC.2021.3088681>
  - [54] M. Mozaffari, X. Lin, and S. Hayes, "Toward 6G with connected sky: UAVs and beyond," *IEEE Communications Magazine*, vol. 59, no. 12, pp. 74–80, 2021.
  - [55] G. Geraci, A. Garcia-Rodriguez, M. M. Azari, A. Lozano, M. Mezzavilla, S. Chatzinotas, Y. Chen, S. Rangan, and M. D. Renzo, "What will the future of UAV cellular communications be? A flight from 5G to 6G," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1304–1335, 2022.
  - [56] J. Luo, Z. Wang, M. Xia, L. Wu, Y. Tian, and Y. Chen, "Path planning for UAV communication networks: Related technologies, solutions, and opportunities," *ACM Comput. Surv.*, vol. 55, no. 9, jan 2023. [Online]. Available: <https://doi.org/10.1145/3560261>
  - [57] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, 1948.
  - [58] A. J. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.
  - [59] M. Malmirchegini and Y. Mostofi, "On the spatial predictability of communication channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 964–978, March 2012.
  - [60] Qualcomm, "Lte unmanned aircraft systems: Trial report (v1.0.1)," Study V 1.0.1, 2017.
  - [61] ITU, "Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 to 60 GHz," International Telecommunication Union, Recommendation P.1410-5, 2013.
  - [62] 3GPP, "Technical specification group radio access network; study on enhanced lte support for aerial vehicles (release 15)," Study TR 36.777, 2018.
  - [63] —, "Technical specification group radio access network; study on 3d channel model for lte (release 12)," Study TR 36.777, 2018.
  - [64] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.
  - [65] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *2014 IEEE Global Communications Conference*, 2014, pp. 2898–2904.
  - [66] C. Yan, L. Fu, J. Zhang, and J. Wang, "A comprehensive survey on UAV communication channel modeling," *IEEE Access*, vol. 7, pp. 107 769–107 792, 2019.
  - [67] W. Khawaja, I. Guvenc, D. W. Matolak, U.-C. Fiebig, and N. Schneckenburger, "A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2361–2391, 2019.
  - [68] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2018.
  - [69] C. Han, Y. Wang, Y. Li, Y. Chen, N. A. Abbasi, T. Kürner, and A. F. Molisch, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *Commun. Surveys Tuts.*, vol. 24, no. 3, p. 1670–1707, jul 2022. [Online]. Available: <https://doi.org/10.1109/COMST.2022.3182539>

- [70] H. Kim and J. Choi, "Channel estimation for spatially/temporally correlated massive mimo systems with one-bit adcs," *J. Wireless Commun. Netw.*, vol. 59, no. 5, pp. 2253–2265, 2019.
- [71] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. J. Paulraj, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, 2010.
- [72] M. Torun, H. Cai, and Y. Mostofi, "Spatial prediction of channel signal strength map using deep fully convolutional neural network," in *2022 56th Asilomar Conference on Signals, Systems, and Computers*, 2022, pp. 553–558.
- [73] C. R. Karanam and Y. Mostofi, "A foundation for wireless channel prediction and full ray makeup estimation using an unmanned vehicle," *IEEE Sensors Journal*, vol. 23, no. 18, pp. 21 452–21 462, 2023.
- [74] S. Evmorfos, K. I. Diamantaras, and A. P. Petropulu, "Reinforcement learning for motion policies in mobile relaying networks," *IEEE Transactions on Signal Processing*, vol. 70, pp. 850–861, 2022.
- [75] D. S. Kalogerias and A. P. Petropulu, "Spatially controlled relay beamforming," *IEEE Transactions on Signal Processing*, vol. 66, no. 24, pp. 6418–6433, 2018.
- [76] K. Diamantaras and A. P. Petropulu, "Optimal mobile relay beamforming via reinforcement learning," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2019, pp. 1–6.
- [77] Remcom, "Wireless InSite." [Online]. Available: <https://www.remcom.com/wireless-insite-em-propagation-software>
- [78] Qualcomm, "Snapdragon 8 gen 2 mobile platform," Product Brief, 2023. [Online]. Available: <https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/Snapdragon-8-Gen-2-Product-Brief.pdf>
- [79] "Mostofi-lab wireless channel simulator," <http://dx.doi.org/10.21229/M9WC7V>.
- [80] A. Gonzalez-Ruiz, A. Ghaffarkhah, and Y. Mostofi, "A Comprehensive Overview and Characterization of Wireless Channels for Networked Robotic and Control Systems," *Journal of Robotics*, vol. 5, p. 19, 2011.
- [81] "NYUSIM: Open-source mmwave and sub-THz wireless channel simulator in matlab and ns-3," <https://wireless.engineering.nyu.edu/nyusim-5g-and-6g/>.
- [82] D. S. Kalogerias and A. P. Petropulu, "Spatially controlled relay beamforming: 2-stage optimal policies," *arXiv preprint arXiv:1705.07463*, 2017.
- [83] S. Evmorfos, "Reinforcement learning for motion policies in mobile relaying networks - code implementation," <https://github.com/SpiliosEv/DeepQLearningSIREN>, 2022, software Implementation - Pytorch.
- [84] S. Evmorfos and A. P. Petropulu, "Deep actor-critic for continuous 3d motion control in mobile relay beamforming networks," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5353–5357.
- [85] U. Challita, W. Saad, and C. Bettstetter, "Interference management for cellular-connected UAVs: A deep reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2125–2140, 2019.
- [86] M. Bloch and J. Barros, "Secrecy capacity," in *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011, p. 49–111.
- [87] C. You and R. Zhang, "Hybrid offline-online design for UAV-enabled data harvesting in probabilistic los channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 3753–3768, 2020.
- [88] DJI, "Phantom 4 rtk: Specs," Technical Specifications, 2023. [Online]. Available: <https://www.dji.com/phantom-4/info>
- [89] L. Steven M., *Planning Algorithms*. Cambridge University Press, 2006. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=170093&site=ehost-live>
- [90] L. E. Dubins, "On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents," *American Journal of Mathematics*, vol. 79, no. 3, pp. 497–516, 1957. [Online]. Available: <http://www.jstor.org/stable/2372560>
- [91] D. Bonilla Licea, M. Bonilla E., M. Ghogho, and M. Saska, "Energy-efficient fixed-wing UAV relay with considerations of airframe shadowing," *IEEE Communications Letters*, vol. 27, no. 6, pp. 1550–1554, 2023.
- [92] Y. Mei, Y.-H. Lu, Y. Hu, and C. Lee, "Deployment of mobile robots with energy and timing constraints," *IEEE Transactions on Robotics*, vol. 22, no. 3, pp. 507–522, 2006.
- [93] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329–2345, 2019.
- [94] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.
- [95] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 641–656, 2014.
- [96] L. Peng and R. Vidal, "Block coordinate descent on smooth manifolds: Convergence theory and twenty-one examples," 2023.
- [97] J. Lee and S. Leyffer, *Mixed Integer Nonlinear Programming*, ser. The IMA Volumes in Mathematics and its Applications. Springer, 2012.
- [98] A. Del Pia, S. S. Dey, and M. Molinaro, "Mixed-integer quadratic programming is in np," *Mathematical Programming*, pp. 225–240, 2017.
- [99] H. Y. Benson and Ümit Sağlam, "Mixed-integer second-order cone programming: A survey," in *Theory Driven by Influential Applications*, 2013, ch. Chapter 2, pp. 13–36. [Online]. Available: <https://pubsonline.informs.org/doi/abs/10.1287/educ.2013.0115>
- [100] IBM, "IBM ILOG CPLEX Optimization Studio," 2023.
- [101] Moesk, "MOSEK 10.1," 2023. [Online]. Available: <https://www.mosek.com/>
- [102] Gurobi Optimization, "Gurobi Optimizer," 2023. [Online]. Available: <https://www.gurobi.com/>
- [103] P. Toth and D. Vigo, *The Vehicle Routing Problem*, P. Toth and D. Vigo, Eds. Society for Industrial and Applied Mathematics, 2002. [Online]. Available: <https://epubs.siam.org/doi/abs/10.1137/1.9780898718515>
- [104] J. Pasha, Z. Elmi, S. Purkayastha, A. M. Fathollahi-Fard, Y.-E. Ge, Y.-Y. Lau, and M. A. Dulebenets, "The drone scheduling problem: A systematic state-of-the-art review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14 224–14 247, 2022.
- [105] S. H. Chung, B. Sah, and J. Lee, "Optimization for drone and drone-truck combined operations: A review of the state of the art and future directions," *Computers & Operations Research*, vol. 123, p. 105004, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0305054820301210>
- [106] N. A. H. Agatz, P. C. Bouman, and M. Schmidt, "Optimization approaches for the traveling salesman problem with drone," *Economics: Mathematical Methods & Programming eJournal*, 2016.
- [107] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
- [108] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research*, vol. 30, pp. 846 – 894, 2011. [Online]. Available: <https://api.semanticscholar.org/CorpusID:14876957>
- [109] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton Univ. Press, 2011.
- [110] D. Bertsekas, *Dynamic Programming and Optimal Control, 4th ed.* Athena Scientific, 2017, vol. 1.
- [111] J. P. Hespanha, *Linear Systems Theory*. Princeton Univ. Press, 2009.
- [112] Z. Shaikh, S. Baidya, and M. Levorato, "Robust multi-path communications for UAVs in the urban lot," in *2018 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*. IEEE, 2018, pp. 1–5.
- [113] A. Feriani, A. Mezghani, and E. Hossain, "On the robustness of deep reinforcement learning in irs-aided wireless communications systems," *arXiv preprint arXiv:2107.08293*, 2021.
- [114] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [115] D. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific, 2019.
- [116] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.
- [117] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [118] G. Dulac-Arnold, R. Evans, H. van Hasselt, P. Sunehag, T. Lillicrap, J. Hunt, T. Mann, T. Weber, T. Degris, and B. Coppin, "Deep reinforcement learning in large discrete action spaces," *arXiv preprint arXiv:1512.07679*, 2015.
- [119] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*. Pmlr, 2014, pp. 387–395.
- [120] M. Köppen, "The curse of dimensionality," in *5th online world conference on soft computing in industrial applications (WSC5)*, vol. 1, 2000, pp. 4–8.

- [121] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [122] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*. PMLR, 2015, pp. 1889–1897.
- [123] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [124] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.
- [125] T. M. Moerland, J. Broekens, A. Plaat, C. M. Jonker *et al.*, "Model-based reinforcement learning: A survey," *Foundations and Trends® in Machine Learning*, vol. 16, no. 1, pp. 1–118, 2023.
- [126] A. G. Barto, "Intrinsic motivation and reinforcement learning," *Intrinsically motivated learning in natural and artificial systems*, pp. 17–47, 2013.
- [127] S. Evmorfos, K. I. Diamantaras, and A. P. Petropulu, "Reinforcement learning for motion policies in mobile relaying networks," *IEEE Transactions on Signal Processing*, vol. 70, pp. 850–861, 2022.
- [128] K. Diamantaras and A. Petropulu, "Optimal mobile relay beamforming via reinforcement learning," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2019, pp. 1–6.
- [129] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas, "Dueling network architectures for deep reinforcement learning," in *International conference on machine learning*. PMLR, 2016, pp. 1995–2003.
- [130] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *International conference on machine learning*. PMLR, 2018, pp. 1587–1596.
- [131] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [132] H. Kumar, D. S. Kalogerias, G. J. Pappas, and A. Ribeiro, "Zeroth-order deterministic policy gradient," *arXiv preprint arXiv:2006.07314*, 2020.
- [133] A. Alidadi Shamsabadi, A. Yadav, O. Abbasi, and H. Yanikomeroglu, "Handling interference in integrated HAPS-terrestrial networks through radio resource management," *IEEE Wireless Communications Letters*, vol. 11, no. 12, pp. 2585–2589, 2022.
- [134] X. Fang, W. Feng, Y. Wang, Y. Chen, N. Ge, Z. Ding, and H. Zhu, "Noma-based hybrid satellite-uav-terrestrial networks for 6G maritime coverage," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 138–152, 2023.
- [135] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-d placement of an aerial base station in next generation cellular networks," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–5.
- [136] Q. Ding, Y. Luo, R. Yang, J. Hu, and C. Luo, "Joint deployment, beamforming and power allocation of mmwave full-duplex UAV-BS," in *Proceedings of the 5th International ACM Mobicom Workshop on Drone Assisted Wireless Communications for 5G and Beyond*, ser. DroneCom '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 127–132. [Online]. Available: <https://doi.org/10.1145/3555661.3560876>
- [137] M. Mozaffari, A. Taleb Zadeh Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5G with UAVs: Foundations of a 3D wireless cellular network," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 357–372, 2019.
- [138] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-uav enabled wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.
- [139] H. Bayerlein, P. De Kerret, and D. Gesbert, "Trajectory optimization for autonomous flying base station via reinforcement learning," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, pp. 1–5.
- [140] Z. Qin, Z. Liu, G. Han, C. Lin, L. Guo, and L. Xie, "Distributed UAV-BSS trajectory optimization for user-level fair communication service with multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12 290–12 301, 2021.
- [141] C. Zhan and Y. Zeng, "Energy minimization for cellular-connected UAV: From optimization to deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5541–5555, 2022.
- [142] Y. Zeng and X. Xu, "Path design for cellular-connected UAV with reinforcement learning," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.
- [143] S. Yin, S. Zhao, Y. Zhao, and F. R. Yu, "Intelligent trajectory design in UAV-aided communications with reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8227–8231, 2019.
- [144] H. Mohammadi, V. Marojevic, and B. Shang, "Analysis of reinforcement learning schemes for trajectory optimization of an aerial radio unit," *arXiv preprint arXiv:2211.10524*, 2022.
- [145] Y. Li, X. Yuan, Y. Hu, J. Yang, and A. Schmeink, "Optimal UAV trajectory design for moving users in integrated sensing and communications networks," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–18, 2023.
- [146] K. Meng, Q. Wu, S. Ma, W. Chen, K. Wang, and J. Li, "Throughput maximization for UAV-enabled integrated periodic sensing and communication," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 671–687, 2023.
- [147] D. S. Kalogerias and A. P. Petropulu, "Mobile beamforming & spatially controlled relay communications," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 6405–6409.
- [148] Y. Yan and Y. Mostofi, "Robotic router formation in realistic communication environments," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 810–827, July 2012.
- [149] J. Fink, A. Ribeiro, and V. Kumar, "Robust control of mobility and communications in autonomous robot teams," *IEEE Access*, vol. 1, pp. 290–309, 2013.
- [150] A. Muralidharan and Y. Mostofi, "Energy optimal distributed beamforming using unmanned vehicles," *IEEE Transactions on Control of Network Systems*, 2017.
- [151] —, "Distributed beamforming using mobile robots," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 6385–6389.
- [152] S. Evmorfos, K. I. Diamantaras, and A. P. Petropulu, "Reinforcement learning for motion policies in mobile relaying networks," *IEEE Transactions on Signal Processing*, vol. 70, pp. 850–861, 2022.
- [153] S. Evmorfos and A. P. Petropulu, "Deep actor-critic for continuous 3d motion control in mobile relay beamforming networks," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5353–5357.
- [154] D. S. Kalogerias and A. P. Petropulu, "Mobi-cliques for improving ergodic secrecy in fading wiretap channels under power constraints," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 1578–1591.
- [155] S. S. Hassan, Y. M. Park, Y. K. Tun, W. Saad, Z. Han, and C. S. Hong, "3TO: THz-enabled throughput and trajectory optimization of UAVs in 6G networks by proximal policy optimization deep reinforcement learning," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 5712–5718.
- [156] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *2015 IEEE global communications conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.
- [157] M. J. Farooq and Q. Zhu, "A multi-layer feedback system approach to resilient connectivity of remotely deployed mobile internet of things," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 2, pp. 422–432, 2018.
- [158] Z. Kang, C. You, and R. Zhang, "3d placement for multi-uav relaying: An iterative gibbs-sampling and block coordinate descent optimization approach," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 2047–2062, 2020.
- [159] S. Y. Shin *et al.*, "Sub-connected hybrid precoding and trajectory optimization using deep reinforcement learning for energy-efficient millimeter-wave UAV communications," *IEEE Wireless Communications Letters*, 2023.
- [160] C. Villani, *Optimal Transport: Old and New*. Springer, 2008.
- [161] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [162] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1295–1315, 2021.
- [163] K. Meng, Q. Wu, J. Xu, W. Chen, Z. Feng, R. Schober, and A. L. Swindlehurst, "UAV-enabled integrated sensing and communication:

- Opportunities and challenges,” *IEEE Wireless Communications*, pp. 1–9, 2023.
- [164] J. Mu, R. Zhang, Y. Cui, N. Gao, and X. Jing, “UAV meets integrated sensing and communication: Challenges and future directions,” *IEEE Communications Magazine*, vol. 61, no. 5, pp. 62–67, 2023.
- [165] C. Chen, H. Song, Q. Li, F. Meneghello, F. Restuccia, and C. Cordeiro, “Wi-Fi sensing based on IEEE 802.11bf,” *IEEE Communications Magazine*, vol. 61, no. 1, pp. 121–127, 2023.
- [166] I. Griva and R. J. Vanderbei, “Case studies in optimization: Catenary problem,” *Optimization and Engineering*, vol. 6, pp. 463–482, 2005. [Online]. Available: <https://vanderbei.princeton.edu/tex/chain/chain.pdf>
- [167] B. Zafar, S. Gherekhloo, and M. Haardt, “Analysis of multihop relaying networks: Communication between range-limited and cooperative nodes,” *IEEE Vehicular Technology Magazine*, vol. 7, no. 3, pp. 40–47, 2012.
- [168] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [169] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, “Control barrier functions: Theory and applications,” in *2019 18th European Control Conference (ECC)*, 2019, pp. 3420–3431.
- [170] S. J. Wright, “Coordinate descent algorithms,” *Mathematical programming*, vol. 151, no. 1, pp. 3–34, 2015.
- [171] A. E. Gelfand, “Gibbs sampling,” *Journal of the American statistical Association*, vol. 95, no. 452, pp. 1300–1304, 2000.
- [172] G. Zheng, K.-K. Wong, A. Paulraj, and B. Ottersten, “Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation,” *IEEE Signal Processing Letters*, vol. 16, no. 4, pp. 257–260, 2009.
- [173] D. Pisinger, “A minimal algorithm for the multiple-choice knapsack problem,” *European Journal of Operational Research*, vol. 83, no. 2, pp. 394–410, 1995, eURO Summer Institute Combinatorial Optimization. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0377221795000151>
- [174] E. Koyuncu, Y. Jing, and H. Jafarkhani, “Distributed beamforming in wireless relay networks with quantized feedback,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1429–1439, 2008.
- [175] M. Sabin and R. Gray, “Global convergence and empirical consistency of the generalized lloyd algorithm,” *IEEE Transactions on information theory*, vol. 32, no. 2, pp. 148–155, 1986.
- [176] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, “Distributed beamforming for relay networks based on second-order statistics of the channel state information,” *IEEE transactions on signal processing*, vol. 56, no. 9, pp. 4306–4316, 2008.
- [177] D. S. Kalogerias, N. Chatzipanagiotis, M. M. Zavlanos, and A. P. Petropulu, “Mobile jammers for secrecy rate maximization in cooperative networks,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 2901–2905.
- [178] W. Hurst, H. Cai, and Y. Mostofi, “Communication-aware RRT\*: Path planning for robotic communication operation in obstacle environments,” in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [179] S. Zhang, Y. Zeng, and R. Zhang, “Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective,” *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2580–2604, 2019.
- [180] N. Cherif, W. Jaafar, H. Yanikomeroglu, and A. Yongacoglu, “Disconnectivity-aware energy-efficient cargo-uav trajectory planning with minimum handoffs,” in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [181] A. Muralidharan and Y. Mostofi, “Path planning for minimizing the expected cost until success,” *IEEE Transactions on Robotics*, vol. 35, no. 2, pp. 466–481, 2019.
- [182] —, “Statistics of the Distance Traveled until Connectivity for Unmanned Vehicles,” *Autonomous Robots*, vol. 44, no. 1, pp. 25–42, 2020.
- [183] Y. Yan and Y. Mostofi, “Co-optimization of communication and motion planning of a robotic operation under resource constraints and in fading environments,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1562–1572, April 2013.
- [184] U. Ali, Y. Yan, Y. Mostofi, and Y. Wardi, “An optimal control approach for communication and motion co-optimization in realistic fading environment.”
- [185] U. Ali, H. Cai, Y. Mostofi, and Y. Wardi, “Motion-communication co-optimization with cooperative load transfer in mobile robotics: An optimal control perspective,” *IEEE Transactions on Control of Network Systems*, vol. 6, no. 2, pp. 621–632, 2019.
- [186] K. K. Nguyen, T. Q. Duong, T. Do-Duy, H. Claussen, and L. Hanzo, “3d UAV trajectory and data collection optimisation via deep reinforcement learning,” *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2358–2371, 2022.
- [187] W. Hurst and Y. Mostofi, “Optimization of mobile robotic relay operation for minimal average wait time,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 6, pp. 3733–3747, 2023.
- [188] —, “Optimal dynamic trajectories for UAVs in mobility-enabled relay systems,” in *Proceedings of the 62nd IEEE Conference on Decision and Control (CDC)*, 2023, To appear.
- [189] O. S. Oubbati, M. Atiquzzaman, H. Lim, A. Rachedi, and A. Lakas, “Synchronizing UAV teams for timely data collection and energy transfer by deep reinforcement learning,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6682–6697, 2022.
- [190] K. Li, W. Ni, E. Tovar, and M. Guizani, “Joint flight cruise control and data collection in UAV-aided internet of things: An onboard deep reinforcement learning approach,” *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9787–9799, 2021.
- [191] Y. Yan and Y. Mostofi, “To go or not to go on energy-aware and communication-aware robotic operation,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 218 – 231, July 2014.
- [192] H. Cai and Y. Mostofi, “Co-optimization of motion, communication, and sensing in real wireless channel environments via monte carlo tree search,” *IEEE Transactions on Control of Network Systems*, vol. 9, no. 3, pp. 1493–1505, 2022.
- [193] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, 2017.
- [194] M. Samir, S. Sharafeddine, C. M. Assi, T. M. Nguyen, and A. Ghayeb, “UAV trajectory planning for data collection from time-constrained IoT devices,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 34–46, 2020.
- [195] O. Ghdiri, W. Jaafar, S. Alfattani, J. B. Abderrazak, and H. Yanikomeroglu, “Energy-efficient multi-uav data collection for IoT networks with time deadlines,” in *Proc. IEEE IEEE Global Telecommun. Conf. (GLOBECOM)*, 2020, pp. 1–6.
- [196] A. Ghaffarkhah and Y. Mostofi, “Dynamic networked coverage of time-varying environments in the presence of fading communication channels,” *ACM Transactions on Sensor Networks*, vol. 10, no. 3, April 2014.
- [197] X. Gao, X. Zhu, and L. Zhai, “AoI-sensitive data collection in multi-UAV-assisted wireless sensor networks,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5185–5197, 2023.
- [198] M. A. Abd-Elmagid and H. S. Dhillon, “Average peak age-of-information minimization in UAV-assisted IoT networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 2003–2008, 2019.
- [199] R. J. Kerzowski and J. H. Griner, “Control and non-payload communications links for integrated unmanned aircraft operations,” in *2012 Joint Conference - 18th Ka and Broadband*, 2012.
- [200] L. M. Ricciardi and S. Sato, “First-passage-time density and moments of the Ornstein-Uhlenbeck process,” *Journal of Applied Probability*, vol. 25, no. 1, pp. 43–57, 1988.
- [201] E. F. Beckenbach and R. Weller, “Problem of bolza,” in *Modern Mathematics for the Engineer - First Series*. Dover Publications, 2013, ch. 4.12. [Online]. Available: <https://app.knovel.com/hotlink/khtml/id:kt011YUJF1/modern-mathematics-engineer/problem-of-bolza>
- [202] R. Shah, S. Roy, S. Jain, and W. Brunette, “Data mules: modeling a three-tier architecture for sparse sensor networks,” in *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications, 2003.*, 2003, pp. 30–41.
- [203] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksall, and N. B. Shroff, “Update or wait: How to keep your data fresh,” *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.
- [204] V. Vishnevsky and O. Semenova, “Polling systems and their application to telecommunication networks,” *Mathematics*, vol. 9, no. 2, 2021.
- [205] M. Ahmed, R. Seraj, and S. M. S. Islam, “The k-means algorithm: A comprehensive survey and performance evaluation,” *Electronics*, vol. 9, no. 8, 2020. [Online]. Available: <https://www.mdpi.com/2079-9292/9/8/1295>
- [206] C. Blum, “Ant colony optimization: Introduction and recent trends,” *Physics of Life Reviews*, vol. 2, no. 4, pp. 353–373, 2005. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1571064505000333>
- [207] R. L. Disney and D. König, “Queueing networks: A survey of their random processes,” *SIAM Review*, vol. 27, no. 3, pp. 335–403, 1985. [Online]. Available: <http://www.jstor.org/stable/2031581>

- [208] C. J. Watkins and P. Dayan, “Q-learning,” *Machine learning*, vol. 8, pp. 279–292, 1992.
- [209] S. K. Singh, R. Singh, and B. Kumbhani, “The evolution of radio access network towards open-ran: Challenges and opportunities,” in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2020, pp. 1–6.
- [210] D. Zhao, H. Wang, K. Shao, and Y. Zhu, “Deep reinforcement learning with experience replay based on sarsa,” in *2016 IEEE symposium series on computational intelligence (SSCI)*. IEEE, 2016, pp. 1–6.
- [211] I. Kostrikov, A. Nair, and S. Levine, “Offline reinforcement learning with implicit q-learning,” *arXiv preprint arXiv:2110.06169*, 2021.
- [212] G. Tesauro *et al.*, “Temporal difference learning and td-gammon,” *Communications of the ACM*, vol. 38, no. 3, pp. 58–68, 1995.
- [213] A. A. Sherstov and P. Stone, “Function approximation via tile coding: Automating parameter choice,” in *International symposium on abstraction, reformulation, and approximation*. Springer, 2005, pp. 194–205.
- [214] Y. Zeng, R. Zhang, and T. J. Lim, “Throughput maximization for UAV-enabled mobile relaying systems,” *IEEE Transactions on communications*, vol. 64, no. 12, pp. 4983–4996, 2016.
- [215] Q. Wang, A. Kakkavas, X. Gong, and R. A. Stirling-Gallacher, “Towards integrated sensing and communications for 6G,” in *2022 2nd IEEE International Symposium on Joint Communications & Sensing (JC&S)*. IEEE, 2022, pp. 1–6.
- [216] J. Tan and L. Dai, “Delay-phase precoding for THz massive MIMO with beam split,” in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.
- [217] S. Evmorfos, K. Diamantaras, and A. Petropulu, “Double deep q learning with gradient biasing for mobile relay beamforming networks,” in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 742–746.
- [218] —, “Deep q learning with fourier feature mapping for mobile relay beamforming networks,” in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 126–130.
- [219] S. Evmorfos, D. Kalogieras, and A. Petropulu, “Adaptive discrete motion control for mobile relay networks,” *Frontiers in Signal Processing*, vol. 2, p. 867388, 2022.
- [220] S. M. Abohashish, R. Y. Rizk, and E. Elsedimy, “Trajectory optimization for UAV-assisted relay over 5g networks based on reinforcement learning framework,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2023, no. 1, p. 55, 2023.
- [221] C. Han, W. Gao, N. Yang, and J. M. Jornet, “Molecular absorption effect: A double-edged sword of terahertz communications,” *IEEE Wireless Communications*, 2022.
- [222] 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, 2019.
- [223] M. T. Spaan, “Partially observable markov decision processes,” in *Reinforcement learning: State-of-the-art*. Springer, 2012, pp. 387–414.
- [224] G. Vasconcelos, R. S. Miani, V. C. Guizilini, and J. R. Souza, “Evaluation of dos attacks on commercial wi-fi-based UAVs,” *International Journal of Communication Networks and Information Security*, vol. 11, no. 1, pp. 212–223, 2019.
- [225] V. Desnitsky and I. Kottenko, “Simulation and assessment of battery depletion attacks on unmanned aerial vehicles for crisis management infrastructures,” *Simulation Modelling Practice and Theory*, vol. 107, p. 102244, 2021.
- [226] D. He, S. Chan, and M. Guizani, “Communication security of unmanned aerial vehicles,” *IEEE Wireless Communications*, vol. 24, no. 4, pp. 134–139, 2016.
- [227] A. D. Wyner, “The wire-tap channel,” *Bell system technical journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [228] S. A. A. Fakoorian and A. L. Swindlehurst, “Solutions for the mimo gaussian wiretap channel with a cooperative jammer,” *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 5013–5022, 2011.
- [229] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, “Improving wireless physical layer security via cooperating relays,” *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1875–1888, 2010.
- [230] G. Zheng, L.-C. Choo, and K.-K. Wong, “Optimal cooperative jamming to enhance physical layer security using relays,” *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1317–1322, 2011.
- [231] J. Li, A. P. Petropulu, and S. Weber, “On cooperative relaying schemes for wireless physical layer security,” *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4985–4997, 2011.
- [232] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, “Improving physical layer secrecy using full-duplex jamming receivers,” *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4962–4974, 2013.
- [233] A. Khisti and G. W. Wornell, “Secure transmission with multiple antennas—part ii: The mimome wiretap channel,” *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, 2010.
- [234] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, 2008.
- [235] N. Su, F. Liu, C. Masouros, T. Ratnarajah, and A. Petropulu, “Secure dual-functional radar-communication transmission: Hardware-efficient design,” in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 629–633.
- [236] N. Su, F. Liu, Z. Wei, Y.-F. Liu, and C. Masouros, “Secure dual-functional radar-communication transmission: Exploiting interference for resilience against target eavesdropping,” *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7238–7252, 2022.
- [237] Y.-K. Li and A. Petropulu, “An irs-assisted secure dual-function radar-communication system,” *arXiv preprint arXiv:2310.00555*, 2023.
- [238] P. McEnroe, S. Wang, and M. Liyanage, “A survey on the convergence of edge computing and ai for UAVs: Opportunities and challenges,” *IEEE Internet of Things Journal*, vol. 9, no. 17, pp. 15435–15459, 2022.
- [239] Z. Liu, Y. Cao, P. Gao, X. Hua, D. Zhang, and T. Jiang, “Multi-uav network assisted intelligent edge computing: Challenges and opportunities,” *China Communications*, vol. 19, no. 3, pp. 258–278, 2022.
- [240] J. Li, Y. Bi, M. Lan, H. Qin, M. Shan, F. Lin, and B. M. Chen, “Real-time simultaneous localization and mapping for UAV: A survey,” in *Proc. of International micro air vehicle competition and conference*, vol. 2016, 2016, p. 237.
- [241] A. Sufian, C. You, and M. Dong, “A deep transfer learning-based edge computing method for home health monitoring,” in *2021 55th annual conference on information sciences and systems (CISS)*. IEEE, 2021, pp. 1–6.
- [242] I. Jang, H. Kim, D. Lee, Y.-S. Son, and S. Kim, “Knowledge transfer for on-device deep reinforcement learning in resource constrained edge computing systems,” *IEEE Access*, vol. 8, pp. 146588–146597, 2020.
- [243] P. S. Chandakkar, Y. Li, P. L. K. Ding, and B. Li, “Strategies for re-training a pruned neural network in an edge computing paradigm,” in *2017 IEEE International Conference on Edge Computing (EDGE)*. IEEE, 2017, pp. 244–247.
- [244] S. Zhu, L. Gui, D. Zhao, N. Cheng, Q. Zhang, and X. Lang, “Learning-based computation offloading approaches in UAVs-assisted edge computing,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 928–944, 2021.
- [245] Y. Cheng, J. Lu, D. Niyato, B. Lyu, J. Kang, and S. Zhu, “Federated transfer learning with client selection for intrusion detection in mobile edge computing,” *IEEE Communications Letters*, vol. 26, no. 3, pp. 552–556, 2022.
- [246] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [247] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, “A survey on federated learning,” *Knowledge-Based Systems*, vol. 216, p. 106775, 2021.
- [248] Y. Qu, H. Dai, Y. Zhuang, J. Chen, C. Dong, F. Wu, and S. Guo, “Decentralized federated learning for UAV networks: Architecture, challenges, and opportunities,” *IEEE Network*, vol. 35, no. 6, pp. 156–162, 2021.
- [249] H. Zhang and L. Hanzo, “Federated learning assisted multi-uav networks,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 14104–14109, 2020.
- [250] Y. Wang, Z. Su, N. Zhang, and A. Benslimane, “Learning in the air: Secure federated learning for UAV-assisted crowdsensing,” *IEEE Transactions on network science and engineering*, vol. 8, no. 2, pp. 1055–1069, 2020.